Adaptive estimator of the memory parameter and goodness-of-fit test using a multidimensional increment ratio statistic

Jean-Marc Bardet, Béchir Dola

Adaptive estimator of the memory parameter and goodness-of-fit test using a multidimensional increment ratio statistic

Jean-Marc Bardet and Béchir Dola

bardet@univ-paris1.fr , bechir.dola@malix.univ-paris1.fr

SAMM, Université Panthéon-Sorbonne (Paris I), 90 rue de Tolbiac, 75013 Paris, FRANCE

September 22, 2011

Abstract

The increment ratio (IR) statistic was first defined and studied in Surgailis *et al.* (2007) for estimating the memory parameter either of a stationary or an increment stationary Gaussian process. Here three extensions are proposed in the case of stationary processes. Firstly, a multidimensional central limit theorem is established for a vector composed by several IR statistics. Secondly, a goodness-of-fit χ^2-type test can be deduced from this theorem. Finally, this theorem allows to construct adaptive versions of the estimator and test which are studied in a general semiparametric frame. The adaptive estimator of the long-memory parameter is proved to follow an oracle property. Simulations attest of the interesting accuracies and robustness of the estimator and test, even in the non Gaussian case.

Keywords: Long-memory Gaussian processes; goodness-of-fit test; estimation of the memory parameter; minimax adaptive estimator.

1 Introduction

After almost thirty years of intensive and numerous studies, the long-memory processes form now an important topic of the time series study (see for instance the book edited by Doukhan *et al*, 2003). The most famous long-memory stationary time series are the fractional Gaussian noises (fGn) with Hurst parameter H and FARIMA(p, d, q) processes. For both these time series, the spectral density f in 0 follows a power law: $f(\lambda) \sim C\,\lambda^{-2d}$ where $H = d + 1/2$ in the case of the fGn. In the case of long memory process $d \in (0, 1/2)$ but a natural expansion to $d \in (-1/2, 0]$ (short memory) implied that d can be considered more generally as a memory parameter.

There are a lot of statistical results relative to the estimation of this memory parameter d. First and main results in this direction have been obtained for parametric models with the essential articles of Fox and Taqqu (1986) and Dahlhaus (1989) for Gaussian time series, Giraitis and Surgailis (1990) for linear processes and Giraitis and Taqqu (1999) for non linear functions of Gaussian processes.

However parametric estimators are not really robust and can induce no consistent estimations. Thus, the research is now rather focused on semiparametric estimators of the memory parameter. Different approaches were considered: the famous R/S statistic (see Hurst, 1951), the log-periodogram estimator (studied firstly by Geweke and Porter-Hudack, 1983, notably improved by Robinson, 1995a, and Moulines and Soulier, 2003), the local Whittle estimator (see Robinson, 1995b) or the wavelet based estimator (see Veitch *et al*, 2003, Moulines *et al*, 2007 or Bardet *et al*, 2008). All these estimators require the choice of an auxiliary parameter (frequency bandwidth, scales, etc.) but adaptive versions of these estimators are generally built for avoiding this choice. In a general semiparametric frame, Giraitis *et al* (1997) obtained the asymptotic lower bound for the minimax risk in the estimation of d, expressed as a function of the second order parameter of the spectral density expansion around 0. Several adaptive semiparametric estimators are proved to follow an oracle property up to multiplicative logarithm term. But simulations (see for instance Bardet *et al*, 2003 or 2008) show that the most accurate estimators are local Whittle, global log-periodogram and wavelet based estimators.

In this paper, we consider the IR (Increment Ratio) estimator of long-memory parameter (see its definition in the next Section) for Gaussian time series recently introduced in Surgailis *et al.* (2007) and we propose three extensions. Firstly, a multivariate central limit theorem is established for a vector of IR statistics with different "windows" (see Section 2) and this induces to consider a pseudo-generalized least square estimator of the parameter d. Secondly, this multivariate result allows us to define an adaptive estimator of the memory parameter d based on IR statistics: an "optimal" window is automatically computed (see Section 3). This notably improves the results of Surgailis *et al.* (2007) in which the choice of m is either theoretical (and cannot be applied to data) or guided by empirical rules without justifications. Thirdly, an adaptive goodness-of-fit test is deduced and its convergence to a chi-square distribution is established (see Section 3).
In Section 4, several Monte Carlo simulations are realized for optimizing the adaptive estimator and exhibiting the theoretical results. Then some numerical comparisons are made with the 3 semiparametric estimators previously mentioned (local Whittle, global log-periodogram and wavelet based estimators) and the results are even better than the theory seems to indicate: as well in terms of convergence rate than in terms of robustness (notably in case of trend or seasonal component), the adaptive IR estimator and goodness-of-fit test provide efficient results. Finally, all the proofs are grouped in Section 5.

2 The multidimensional increment ratio statistic and its statistical applications

Let $X = (X_k)_{k\in\mathbb{N}}$ be a Gaussian time series satisfying the following Assumption $S(d,\beta)$:

Assumption $S(d,\beta)$**:** *There exist $\varepsilon > 0$, $c_0 > 0$, $c_0' > 0$ and $c_1 \in \mathbb{R}$ such that $X = (X_t)_{t\in\mathbb{Z}}$ is a stationary Gaussian time series having a spectral density f satisfying for all $\lambda \in (-\pi,0)\cup(0,\pi)$*

$$f(\lambda) = c_0|\lambda|^{-2d} + c_1|\lambda|^{-2d+\beta} + O\big(|\lambda|^{-2d+\beta+\varepsilon}\big) \quad \text{and} \quad |f'(\lambda)| \le c_0'\,\lambda^{-2d-1}. \tag{2.1}$$

Remark 1. *Note that here we only consider the case of stationary processes. However, as it was already done in Surgailis* et al. *(2007), it could be possible,* mutatis mutandis*, to extend our results to the case of processes having stationary increments.*

Let $(X_1,\cdots,X_N)$ be a path of X. For $m\in\mathbb{N}^*$, define the random variable $IR_N(m)$ such as

$$IR_N(m) := \frac{1}{N-3m}\sum_{k=0}^{N-3m-1}\frac{\big|\big(\sum_{t=k+1}^{k+m}X_{t+m}-\sum_{t=k+1}^{k+m}X_t\big)+\big(\sum_{t=k+m+1}^{k+2m}X_{t+m}-\sum_{t=k+m+1}^{k+2m}X_t\big)\big|}{\big|\big(\sum_{t=k+1}^{k+m}X_{t+m}-\sum_{t=k+1}^{k+m}X_t\big)\big|+\big|\big(\sum_{t=k+m+1}^{k+2m}X_{t+m}-\sum_{t=k+m+1}^{k+2m}X_t\big)\big|}.$$

From Surgailis *et al.* (2007), with m such that $N/m\to\infty$ and $m\to\infty$,

$$\sqrt{\frac{N}{m}}\big(IR_N(m)-\mathrm{E}IR_N(m)\big) \xrightarrow[N\to\infty]{\mathcal{L}} \mathcal{N}(0,\sigma^2(d)),$$

where

$$\sigma^2(d) \;:=\; 2\int_0^\infty \mathrm{Cov}\Big(\frac{|Z_d(0)+Z_d(1)|}{|Z_d(0)|+|Z_d(1)|}\,,\,\frac{|Z_d(\tau)+Z_d(\tau+1)|}{|Z_d(\tau)|+|Z_d(\tau+1)|}\Big)d\tau \tag{2.2}$$

$$\text{and}\quad Z_d(\tau) \;:=\; \frac{1}{\sqrt{|4^{d+0.5}-4|}}\big(B_{d+0.5}(\tau+2)-2\,B_{d+0.5}(\tau+1)+B_{d+0.5}(\tau)\big) \tag{2.3}$$

with B_H a standardized fractional Brownian motion (FBM) with Hurst parameter $H \in (0,1)$.

Remark 2. *This convergence was obtained for Gaussian processes in Surgailis et al. (2007), but there also exist results concerning a modified IR statistic applied to stable processes (see Vaiciulis, 2009) with a different kind of limit theorem. We may suspect that it is also possible to extend the previous central limit theorem to long memory linear processes (since a Donsker type theorem with FBM as limit was proved for long memory linear processes, see for instance Ho and Hsing, 1997) but such result requires to prove a non obvious central limit theorem for a functional of multidimensional linear process. Surgailis et al. (2007) also considered the case of i.i.d.r.v. in the domain of attraction of a stable law with index $0<\alpha<2$ and skewness parameter $-1\leq\beta\leq 1$ and concluded that $IR_N(m)$ converges to almost the same limit. Finally, in Bardet and Surgailis (2011) a "continuous" version of the IR statistic is considered for several kind of continuous time processes (Gaussian processes, diffusions and Lévy processes).*

Now, instead of this univariate IR statistic, define a multivariate IR statistic as follows: let $m_j = j\,m,\ j = 1,\cdots,p$ with $2 \leq p[N/m]-4$, and define the random vector $(IR_N(j\,m))_{1\leq j\leq p}$. Thus, p is the number of considered window lengths of this multivariate statistic. In the sequel we naturally extend the results obtained for $m\in\mathbb{N}^*$ to $m\in(0,\infty)$ by the convention: $(IR_N(j\,m))_{1\leq j\leq p} = (IR_N(j\,[m]))_{1\leq j\leq p}$ (which change nothing to the asymptotic results).
We can establish a multidimensional central limit theorem satisfied by $(IR_N(j\,m))_{1\leq j\leq p}$:

Property 2.1. *Assume that Assumption $S(d,\beta)$ holds with $-0.5<d<0.5$ and $\beta>0$. Then*

$$\sqrt{\frac{N}{m}}\Big(IR_N(j\,m)-\mathrm{E}\big[IR_N(j\,m)\big]\Big)_{1\leq j\leq p} \xrightarrow[{[N/m]\wedge m\to\infty}]{\mathcal{L}} \mathcal{N}(0,\Gamma_p(d)) \tag{2.4}$$

with $\Gamma_p(d) = (\sigma_{i,j}(d))_{1\leq i,j\leq p}$ where for $t\in\mathbb{R}$

$$\sigma_{i,j}(d) := \int_{-\infty}^{\infty}\mathrm{Cov}\Big(\frac{|Z_d^{(i)}(0)+Z_d^{(i)}(i)|}{|Z_d^{(i)}(0)|+|Z_d^{(i)}(i)|}\,,\,\frac{|Z_d^{(j)}(\tau)+Z_d^{(j)}(\tau+j)|}{|Z_d^{(j)}(\tau)|+|Z_d^{(j)}(\tau+j)|}\Big)d\tau$$

$$\textit{and}\quad Z_d^{(j)}(\tau) = \frac{1}{\sqrt{|4^{d+0.5}-4|}}\big(B_{d+0.5}(\tau+2j)-2B_{d+0.5}(\tau+j)+B_{d+0.5}(\tau)\big). \tag{2.5}$$

The proof of this property as well as all the other proofs are given in Appendix. Moreover we will assume in the sequel that $\Gamma_p(d)$ is a definite positive matrix for all $d\in(-0.5,0.5)$.

Remark 3. *Note that Assumption $S(d,\beta)$ are a little stronger than the conditions required in Surgailis et al. (2007) where f is supposed to satisfy $f(\lambda) = c_0|\lambda|^{-2d}+O(|\lambda|^{-2d+\beta})$ and $|f'(\lambda)|\leq c_0'\,\lambda^{-2d-1}$. Note that Property 2.1 and following Theorem 1 and Proposition 1 are as well checked under these assumptions of Surgailis et al. (2007) even if $\beta\geq 2d+1$ (case which is not consider in their Theorem 2.4). However our automatic procedure for choosing an adaptive scale $\widetilde{m}_N$ requires to specify the second order of the expansion of f and we prefer to already give results under such assumption.*

As in Surgailis *et al.* (2007), for $r\in(-1,1)$, define the function $\Lambda(r)$ by

$$\Lambda(r) := \frac{2}{\pi}\,\arctan\sqrt{\frac{1+r}{1-r}}+\frac{1}{\pi}\sqrt{\frac{1+r}{1-r}}\log(\frac{2}{1+r}). \tag{2.6}$$

and for $d \in (-0.5, 1.5)$ let

$$\Lambda_0(d) \quad := \quad \Lambda(\rho(d)) \quad \text{where} \quad \rho(d) := \frac{4^{d+1.5} - 9^{d+0.5} - 7}{2(4 - 4^{d+0.5})}. \tag{2.7}$$

The function $d \in (-0.5, 1.5) \to \Lambda_0(d)$ is a $\mathcal{C}^\infty$ increasing function. Now, Property 5.1 (see in Section 5) provides the asymptotic behavior of $\mathrm{E}[IR(m)]$ when $m \to \infty$, which is $\mathrm{E}[IR(m)] \sim \Lambda_0(d) + Cm^{-\beta}$ if $\beta < 2d+1$, $\mathrm{E}[IR(m)] \sim \Lambda_0(d) + Cm^{-\beta}\log m$ if $\beta = 2d+1$ and $\mathrm{E}[IR(m)] \sim \Lambda_0(d) + O(m^{-(2d+1)})$ if $\beta > 2d+1$ (C is a non vanishing real number depending on d and β). Therefore by choosing m and N such as $\big(\sqrt{N/m}\big)m^{-\beta} \to 0$, $\big(\sqrt{N/m}\big)m^{-\beta}\log m \to 0$ and $\big(\sqrt{N/m}\big)m^{-(2\beta+1)} \to 0$ (respectively) when $m, N \to \infty$, the term $\mathrm{E}[IR(jm)]$ can be replaced by $\Lambda_0(d)$ in Property 2.1. Then, using the Delta-method with function $(x_i)_{1 \le i \le p} \mapsto (\Lambda_0^{-1}(x_i))_{1 \le i \le p}$, we obtain:

Theorem 1. *Let $\widehat{d}_N(j\,m) := \Lambda_0^{-1}\big(IR_N(j\,m)\big)$ for $1 \le j \le p$. Assume that Assumption $S(d,\beta)$ holds with $-0.5 < d < 0.5$ and $\beta > 0$. Then if $m \sim C\,N^\alpha$ with $C > 0$ and $(1+2\beta)^{-1} \vee (4d+3)^{-1} < \alpha < 1$ then*

$$\sqrt{\frac{N}{m}}\Big(\widehat{d}_N(j\,m) - d\Big)_{1 \le j \le p} \xrightarrow[N\to\infty]{\mathcal{L}} \mathcal{N}\Big(0, (\Lambda_0'(d))^{-2}\,\Gamma_p(d)\Big). \tag{2.8}$$

Remark 4. *If $\beta < 2d+1$, the estimator $\widehat{d}_N(m)$ is a semiparametric estimator of d and its asymptotic mean square error can be minimized with an appropriate sequence (m_N) reaching the well-known minimax rate of convergence for memory parameter d in this semiparametric setting (see for instance Giraitis* et al.*, 1997). Indeed, under Assumption $S(d,\beta)$ with $d \in (-0.5, 0.5)$ and $\beta > 0$ and if $m_N = [N^{1/(1+2\beta)}]$, then the estimator $\widehat{d}_N(m_N)$ is rate optimal in the minimax sense,* i.e.

$$\limsup_{N\to\infty} \sup_{d\in(-0.5,0.5)} \sup_{f\in S(d,\beta)} N^{\frac{2\beta}{1+2\beta}} \cdot \mathrm{E}[(\widehat{d}_N(m_N) - d)^2] < \infty.$$

From the multidimensional CLT (2.8) a pseudo-generalized least square estimation (LSE) of d is possible by defining the following matrix:

$$\widehat{\Sigma}_N(m) := (\Lambda_0'(\widehat{d}_N(m))^{-2}\,\Gamma_p(\widehat{d}_N(m)). \tag{2.9}$$

Since the function $d \in (-0.5, 1.5) \mapsto \sigma(d)/\Lambda'(d)$ is $\mathcal{C}^\infty$ it is obvious that under assumptions of Theorem 1 then

$$\widehat{\Sigma}_N(m) \xrightarrow[N\to\infty]{\mathcal{P}} (\Lambda_0'(d))^{-2}\,\Gamma_p(d).$$

Then with the vector $J_p := (1)_{1 \le j \le p}$ and denoting J_p' its transpose, the pseudo-generalized LSE of d is:

$$\widetilde{d}_N(m) := \big(J_p'\big(\widehat{\Sigma}_N(m)\big)^{-1}J_p\big)^{-1}\,J_p'\,\big(\widehat{\Sigma}_N(m)\big)^{-1}\big(\widehat{d}_N(m_i)\big)_{1 \le i \le p}$$

It is well known (Gauss-Markov Theorem) that the Mean Square Error (MSE) of $\widetilde{d}_N(m)$ is smaller or equal than all the MSEs of $\widehat{d}_N(jm)$, $j = 1, \dots, p$. Hence, we obtain under the assumptions of Theorem 2.8:

$$\sqrt{\frac{N}{m}}\big(\widetilde{d}_N(m) - d\big) \xrightarrow[N\to\infty]{\mathcal{L}} \mathcal{N}\Big(0\,,\, \Lambda_0'(d)^{-2}\,\big(J_p'\,\Gamma_p^{-1}(d)J_p\big)^{-1}\Big), \tag{2.10}$$

and $\Lambda_0'(d)^{-2}\big(J_p'\,\Gamma_p^{-1}(d)J_p\big)^{-1} \le \Lambda_0'(d)^{-2}\sigma^2(d)$.

Now, consider the following test problem: for $(X_1, \cdots, X_n)$ a path of X a Gaussian time series, chose between

- H_0: the spectral density of X satisfies Assumption $S(d,\beta)$ with $-0.5 < d < 0.5$ and $\beta > 0$;
- H_1: the spectral density of X does not satisfy such a behavior.

We deduce from the multidimensional CLT (2.8) a χ^2-type goodness-of-fit test statistic defined by:

$$\widehat{T}_N(m) := \frac{N}{m}\,\big(\widetilde{d}_N(m) - \widehat{d}_N(j\,m)\big)'_{1 \le j \le p}\big(\widehat{\Sigma}_N(m)\big)^{-1}\big(\widetilde{d}_N(m) - \widehat{d}_N(j\,m)\big)_{1 \le j \le p}.$$

Then the following limit theorem can be deduced from Theorem 1:

Proposition 1. *Under the assumptions of Theorem 1 then:*

$$\widehat{T}_N(m) \xrightarrow[N\to\infty]{\mathcal{L}} \chi^2(p-1).$$

3 Adaptive versions of the estimator and goodness-of-fit test

Theorem 1 and Proposition 1 are interesting but they require the knowledge of β to be used (and therefore an appropriated choice of m). We suggest now a procedure (see also Bardet *et al.*, 2008) for obtaining a data-driven selection of an optimal sequence (m_N). For $d \in (-0.5, 1.5)$ and $\alpha \in (0,1)$, define

$$Q_N(\alpha, d) := \big(\widehat{d}_N(j\,N^\alpha) - d\big)'_{1\le j\le p}\big(\widehat{\Sigma}_N(N^\alpha)\big)^{-1}\big(\widehat{d}_N(j\,N^\alpha) - d\big)_{1\le j\le p}. \tag{3.1}$$

Note that by the previous convention, $\widehat{d}_N(j\,N^\alpha) = \widehat{d}_N(j\,[N^\alpha])$ and $\widetilde{d}_N(N^\alpha) = \widetilde{d}_N([N^\alpha])$. Thus $Q_N(\alpha, d)$ corresponds to the sum of the pseudo-generalized squared distance. From previous computations, it is obvious that for a fixed $\alpha \in (0,1)$, Q is minimized by $\widetilde{d}_N(N^\alpha)$ and therefore for $0 < \alpha < 1$ define

$$\widehat{Q}_N(\alpha) := Q_N(\alpha, \widetilde{d}_N(N^\alpha)).$$

It remains to minimize $\widehat{Q}_N(\alpha)$ on $(0,1)$. However, since $\widehat{\alpha}_N$ has to be obtained from numerical computations, the interval $(0,1)$ can be discretized as follows,

$$\widehat{\alpha}_N \in \mathcal{A}_N = \Big\{\frac{2}{\log N}, \frac{3}{\log N}, \ldots, \frac{\log[N/p]}{\log N}\Big\}.$$

Hence, if $\alpha \in \mathcal{A}_N$, it exists $k \in \{2, 3, \ldots, \log[N/p]\}$ such that $k = \alpha \log N$. Consequently, define $\widehat{\alpha}_N$ by

$$\widehat{Q}_N(\widehat{\alpha}_N) := \min_{\alpha\in\mathcal{A}_N} \widehat{Q}_N(\alpha).$$

From the central limit theorem (2.8) one deduces the following :

Proposition 2. *Assume that Assumption $S(d,\beta)$ holds with $-0.5 < d < 0.5$ and $\beta > 0$. Moreover, if $\beta > 2d+1$, suppose that c_0, c_1, c_2, d, β and ε are such that Condition (5.13) or (5.14) holds. Then,*

$$\widehat{\alpha}_N \xrightarrow[N\to\infty]{\mathcal{P}} \alpha^* = \frac{1}{(1+2\beta)\wedge(4d+3)}.$$

Remark 5. *The choice of the set of discretization $\mathcal{A}_N$ is implied by our proof of convergence of $\widehat{\alpha}_N$ to α^*. If the interval $(0,1)$ is stepped in N^c points, with $c > 0$, the used proof cannot attest this convergence. However $\log N$ may be replaced in the previous expression of $\mathcal{A}_N$ by any negligible function of N compared to functions N^c with $c > 0$ (for instance, $(\log N)^a$ or $a \log N$ can be used).*

Remark 6. *The reference to Condition (5.13) or (5.14) is necessary because our proof of the convergence of $\widehat{\alpha}_N$ to α^* requires to know the exact convergence rate of $\mathrm{E}[IR_N(N^\alpha)] - \Lambda_0(d)$ when $\alpha < \alpha^*$. When $\beta \le 2d+1$, since we replaced the conditions on the spectral density of Surgailis et al. (2007) by a second order condition (Assumption $S(d,\beta)$), this convergence rate can be obtained by computations (see Property 5.1). But if $\beta > 2d+1$, we can only obtain $\mathrm{E}[IR_N(N^\alpha)] - \Lambda_0(d) = O(m^{-2d-1})$ under Assumption $S(d,\beta)$: the convergence rate could be slower than m^{-2d-1} and then $\widehat{\alpha}_N$ could converge to $\alpha' < \alpha^*$ (from the proof of Proposition 2). Condition (5.13) and (5.14), which are not very strong, allow to obtain a first order bound for $\mathrm{E}[IR_N(N^\alpha)] - \Lambda_0(d)$ (see Proposition 5.2) and hence to prove $\widehat{\alpha}_N \xrightarrow[N\to\infty]{\mathcal{P}} \alpha^*$.*

From a straightforward application of the proof of Proposition 2, the asymptotic behavior of $\widehat{\alpha}_N$ can be specified, that is,

$$\Pr\Big(\frac{N^{\alpha^*}}{(\log N)^\lambda} \le N^{\widehat{\alpha}_N} \le N^{\alpha^*}\cdot(\log N)^\mu\Big) \xrightarrow[N\to\infty]{} 1, \tag{3.2}$$

for all positive real numbers λ and μ such that $\lambda > \frac{2\alpha^*}{(p-2)(1-\alpha^*)}$ and $\mu > \frac{12}{p-2}$. Consequently, the selected window $\widehat{m}_N = N^{\widehat{\alpha}_N}$ asymptotically growths as N^{α^*} up to a logarithm factor.

Finally, Proposition 2 can be used to define an adaptive estimator of d. First, define the straightforward estimator $\widetilde{d}_N(N^{\widehat{\alpha}_N})$, which should minimize the mean square error using $\widehat{\alpha}_N$. However, the estimator $\widetilde{d}_N(N^{\widehat{\alpha}_N})$ does not satisfy a CLT since $\Pr(\widehat{\alpha}_N \le \alpha^*) > 0$ and therefore it can not be asserted that $\mathrm{E}(\sqrt{N/N^{\widehat{\alpha}_N}}(\widetilde{d}_N(N^{\widehat{\alpha}_N}) - d)) = 0$. To establish a CLT satisfied by an adaptive estimator of d, a (few) shifted sequence of $\widehat{\alpha}_N$, so called $\widetilde{\alpha}_N$, has to be considered to ensure $\Pr(\widetilde{\alpha}_N \le \alpha^*) \underset{N\to\infty}{\longrightarrow} 0$. Hence, consider the adaptive scale sequence $(\widetilde{m}_N)$ such as

$$\widetilde{m}_N := N^{\widetilde{\alpha}_N} \quad \text{with} \quad \widetilde{\alpha}_N := \widehat{\alpha}_N + \frac{6\,\widehat{\alpha}_N}{(p-2)(1-\widehat{\alpha}_N)} \cdot \frac{\log\log N}{\log N}.$$

and the estimator

$$\widetilde{d}_N^{(IR)} := \widetilde{d}_N(\widetilde{m}_N) = \widetilde{d}_N(N^{\widetilde{\alpha}_N}).$$

The following theorem provides the asymptotic behavior of the estimator $\widetilde{d}_N^{(IR)}$:

Theorem 2. *Under assumptions of Proposition 2,*

$$\sqrt{\frac{N}{N^{\widetilde{\alpha}_N}}}\big(\widetilde{d}_N^{(IR)} - d\big) \underset{N\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}\Big(0\,;\,\Lambda_0'(d)^{-2}\left(J_p'\,\Gamma_p^{-1}(d)J_p\right)^{-1}\Big). \tag{3.3}$$

Moreover, if $\beta \le 2d+1$, $\forall \rho > \dfrac{2(1+3\beta)}{(p-2)\beta}$, $\dfrac{N^{\frac{\beta}{1+2\beta}}}{(\log N)^\rho}\cdot\big|\widetilde{d}_N^{(IR)} - d\big| \underset{N\to\infty}{\overset{\mathcal{P}}{\longrightarrow}} 0.$

Remark 7. *When $\beta \le 2d+1$, the adaptive estimator $\widetilde{d}_N^{(IR)}$ converges to d with a rate of convergence rate equal to the minimax rate of convergence $N^{\frac{\beta}{1+2\beta}}$ up to a logarithm factor (this result being classical within this semiparametric framework). Thus there exists $\ell < 0$ such that*

$$N^{\frac{2\beta}{1+2\beta}}(\log N)^\ell\,\mathrm{E}(\widetilde{d}_N^{(IR)} - d)^2 < \infty.$$

Therefore $\widetilde{d}_N^{(IR)}$ satisfies an oracle property for the considered semiparametric model.
If $\beta > 2d+1$, the estimator is not rate optimal. However, simulations (see the following Section) will show that even if $\beta > 2d+1$, the rate of convergence of $\widetilde{d}_N^{(IR)}$ can be better than the one of the best known rate optimal estimators (local Whittle or global log-periodogram estimators).

Moreover an adaptive version of the previous goodness-of-fit test can be derived. Thus define

$$\widetilde{T}_N^{(IR)} := \widehat{T}_N(N^{\widetilde{\alpha}_N}). \tag{3.4}$$

Then,

Proposition 3. *Under assumptions of Proposition 2,*

$$\widetilde{T}_N^{(IR)} \underset{N\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} \chi^2(p-1).$$

4 Simulations and Monte-Carlo experiments

In the sequel, the numerical properties (consistency, robustness, choice of the parameter p) of $\widetilde{d}_N^{(IR)}$ are investigated. Then the simulation results of $\widetilde{d}_N^{(IR)}$ are compared to those obtained with the best known semiparametric long-memory estimators.

Remark 8. *Note that all the softwares (in Matlab language) used in this Section are available with a free access on* `http://samm.univ-paris1.fr/-Jean-Marc-Bardet`.

To begin with, the simulation conditions have to be specified. The results are obtained from 100 generated independent samples of each process belonging to the following "benchmark". The concrete procedures of generation of these processes are obtained from the circulant matrix method, as detailed in Doukhan *et al.* (2003). The simulations are realized for different values of d, N and processes which satisfy Assumption $S(d,\beta)$:

1. the fractional Gaussian noise (fGn) of parameter $H = d + 1/2$ (for $-0.5 < d < 0.5$) and $\sigma^2 = 1$. Such a process is such that Assumption $S(d,2)$ holds;

2. the FARIMA$[p,d,q]$ process with parameter d such that $d \in (-0.5, 0.5)$, the innovation variance σ^2 satisfying $\sigma^2 = 1$ and $p, q \in \mathbb{N}$. A FARIMA$[p,d,q]$ process is such that Assumption $S(d,2)$ holds;

3. the Gaussian stationary process $X^{(d,\beta)}$, such as its spectral density is
$$f_3(\lambda) = \frac{1}{\lambda^{2d}}(1 + \lambda^\beta) \quad \text{for } \lambda \in [-\pi, 0) \cup (0, \pi], \tag{4.1}$$
with $d \in (-0.5, 0.5)$ and $\beta \in (0, \infty)$. Therefore the spectral density f_3 is such as Assumption $S(d,\beta)$ holds.

A "benchmark" which will be considered in the sequel consists of the following particular cases of these processes for $d = -0.4, -0.2, 0, 0.2, 0.4$:

- fGn processes with parameters $H = d + 1/2$;
- FARIMA$[0,d,0]$ processes with standard Gaussian innovations;
- FARIMA$[1,d,1]$ processes with standard Gaussian innovations and AR coefficient $\phi = -0.3$ and MA coefficient $\phi = 0.7$;
- $X^{(d,\beta)}$ Gaussian processes with $\beta = 1$.

4.1 Application of the IR estimator and tests applied to generated data

Choice of the parameter p: This parameter is important to estimate the "beginning" of the linear part of the graph drawn by points $(i, IR(im))_i$. On the one hand, if p is a too small a number (for instance $p = 3$), another small linear part of this graph (even before the "true" beginning N^{α^*}) may be chosen. On the other hand, if p is a too large a number (for instance $p = 50$ for $N = 1000$), the estimator $\widetilde{\alpha}_N$ will certainly satisfy $\widetilde{\alpha}_N < \alpha^*$ since it will not be possible to consider p different windows larger than N^{α^*}. Moreover, it is possible that a "good" choice of p depends on the "flatness" of the spectral density f, *i.e.* on β. We have proceeded to simulations for several values of p (and N and d). Only $\sqrt{MSE}$ of estimators are presented. The results are specified in Table 1.

Conclusions from Table 1: it is clear that $\widetilde{d}_N^{(IR)}$ converges to d for the four processes, the faster for fGn and FARIMA$(0,d,0)$. The optimal choice of p seems to depend on N for the four processes: $\widehat{p} = 10$ for $N = 10^3$, $\widehat{p} = 15$ for $N = 10^4$ and $\widehat{p} \in [15, 20]$ for $N = 10^5$. The flatness of the spectral density of the process does not seem to have any influence, as well as the value of d (result obtained in the detailed simulations). We will adopt in the sequel the choice $\widehat{p} = [1.5 \log(N)]$ reflecting these results. At the contrary to the choice of m, this choice of p only depends on N and even if the adaptive scale $\widetilde{m}_N$ depends on p its value does not change a lot when $p \in \{10, \cdots, 20\}$ for $10^3 \le N \le 10^5$.
Concerning the adaptive choice of m, the main point to be remarked is that the smoother the spectral density the smaller m; thus $\widetilde{m}_N$ is smaller for a trajectory of a fGn or a FARIMA$(0,d,0)$ than for a trajectory of a FARIMA$(1,d,1)$ or $X^{(d,1)}$. The choice of p does not appear to significantly affect the value of $\widetilde{m}_N$. More detailed results show that the larger d included in $(-0.5, 0.5)$ the smaller $\widetilde{m}_N$: for instance, for the fGn, $N = 10^4$

and $p = 15$, the mean of $\widetilde{m}_N$ is respectively equal to 23.9, 8.3, 4.5, 4.2 and 3.8 for d respectively equal to -0.4, -0.2, 0, 0.2 and 0.4. This phenomena can be deduced from the theoretical study since $\alpha^* = (4d+3)^{-1}$ in this case and therefore $\widetilde{m}_N$ almost growths as $N^{(4d+3)^{-1}}$.
Finally, concerning the goodness-of-fit test, we remark that it is too conservative for $p = 5$ or 10 but close to the expected results for $p = 15$ and 20, especially for FARIMA$(1, d, 1)$ or $X^{(d,1)}$.

Asymptotic distributions of the estimator and test: Figure 1 provides the density estimations of $\widetilde{d}_N^{(IR)}$ and $\widetilde{T}_N^{(IR)}$ for 100 independent samples of FGN processes with $d = 0.2$ with $N = 10^4$ for $p = 15$. The goodness-of-fit to the theoretical asymptotic distributions (respectively Gaussian and chi-square) is satisfying.

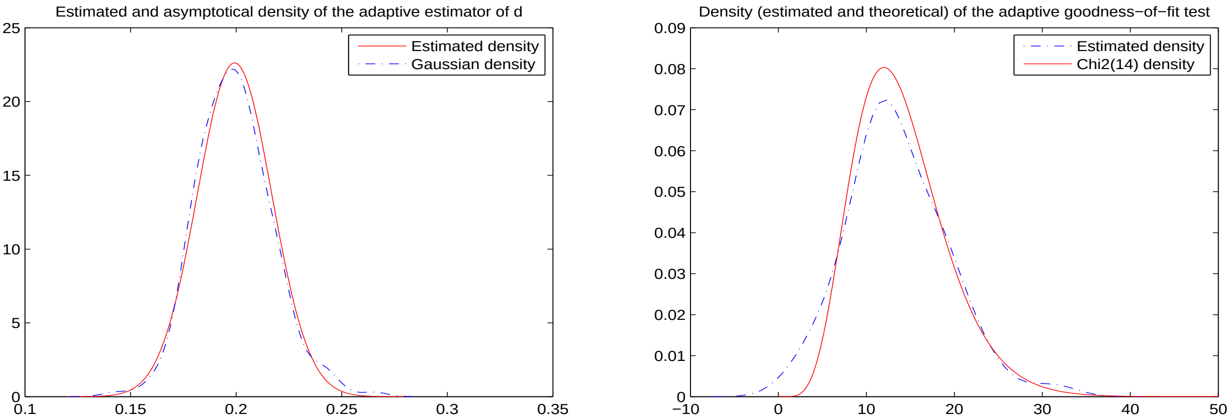

Figure 1: *Density estimations and corresponding theoretical densities of* $\widetilde{d}_N^{(IR)}$ *and* $\widetilde{T}_N^{(IR)}$ *for* 100 *samples of fGn with* $d = 0.2$ *with* $N = 10^4$ *and* $p = 15$.

4.2 Comparison with other adaptive semiparametric estimator of the memory parameter

Consistency of semiparametric estimators: Here we consider the previous "benchmark" and apply the estimator $\widetilde{d}_N^{(IR)}$ and 3 other semiparametric estimators of d known for their accuracies are considered:

- $\widehat{d}_{MS}$ is the adaptive global log-periodogram estimator introduced by Moulines and Soulier (1998, 2003), also called FEXP estimator, with bias-variance balance parameter $\kappa = 2$;
- $\widehat{d}_R$ is the local Whittle estimator introduced by Robinson (1995). The trimming parameter is $m = N/30$;
- $\widehat{d}_W$ is an adaptive wavelet based estimator introduced in Bardet *et al.* (2008) using a Lemarie-Meyer type wavelet (another similar choice could be the adaptive wavelet estimator introduced in Veitch *et al.*, 2003, using a Daubechie's wavelet, but its robustness property are quite less interesting).
- $\widetilde{d}_N^{(IR)}$ defined previously with $p = [1.5 * \log(N)]$.
- $\widehat{d}_N(10)$ and $\widehat{d}_N(30)$ which are the (univariate) IR estimator with $m = 10$ and $m = 30$ respectively, considered in Surgailis *et al.* (2007).

Simulation results are reported in Table 2.

Conclusions from Table 2: The adaptive IR estimator $\widetilde{d}_N^{(IR)}$ numerically shows a convincing convergence rate with respect to the other estimators.
The estimators $\widehat{d}_N(10)$ and $\widehat{d}_N(30)$ are clearly the worst estimators of d. This can be explained by two facts: 1/ the numerical expression of the matrix $\widehat{\Sigma}_N(m)$ is almost a diagonal matrix: therefore a least square regression using several window lengths provides better estimations than an estimator using only one window length; 2/ $\widehat{d}_N(10)$ and $\widehat{d}_N(30)$ use a fixed window length ($m = 10$ and $m = 30$) for any process and N while we know that $m \simeq N^{\alpha^*}$ is the optimal choice which is approximated by $\widetilde{m}_N$.
Both the "spectral" estimator $\widehat{d}_R$ and $\widehat{d}_{MS}$ provide more stable results that do not depend very much on d and the process, while the wavelet based estimator $\widehat{d}_W$ and $\widetilde{d}_N^{(IR)}$ are more sensible to the flatness of the spectral density. But, especially for "smooth processes" (fGn and FARIMA$(0, d, 0)$), $\widetilde{d}_N^{(IR)}$ is a very accurate semiparametric estimator and is globally more efficient than the other estimators.

Robustness of the different semiparametric estimators: To conclude with the numerical properties of the estimators, five different processes not satisfying Assumption $S(d, \beta)$ are considered:

- a FARIMA$(0, d, 0)$ process with innovations satisfying a uniform law;
- a FARIMA$(0, d, 0)$ process with innovations satisfying a symmetric Burr distribution with cumulative distribution function $F(x) = 1-\frac{1}{2}\frac{1}{1+x^2}$ for $x \geq 0$ and $F(x) = \frac{1}{2}\frac{1}{1+x^2}$ for $x \leq 0$ (and therefore $\mathrm{E}|X_i|^2 = \infty$ but $\mathrm{E}|X_i| < \infty$);
- a FARIMA$(0, d, 0)$ process with innovations satisfying a symmetric Burr distribution with cumulative distribution function $F(x) = 1 - \frac{1}{2}\frac{1}{1+|x|^{3/2}}$ for $x \geq 0$ and $F(x) = \frac{1}{2}\frac{1}{1+|x|^{3/2}}$ for $x \leq 0$ (and therefore $\mathrm{E}|X_i|^2 = \infty$ but $\mathrm{E}|X_i| < \infty$);
- a Gaussian stationary process with a spectral density $f(\lambda) = ||\lambda| - \pi/2|^{-2d}$ for all $\lambda \in [-\pi, \pi] \setminus \{-\pi/2, \pi/2\}$: this is a GARMA$(0, d, 0)$ process. The local behavior of f in 0 is $f(|\lambda|) \sim (\pi/2)^{-2d}\,|\lambda|^{-2d}$ with $d = 0$, but the smoothness condition for f in Assumption $S(0, \beta)$ is not satisfied.
- a trended fGn with parameter $H = d + 0.5$ and an additive linear trend;
- a fGn ($H = d + 0.5$) with an additive linear trend and an additive sinusoidal seasonal component of period $T = 12$.

The results of these simulations are given in Table 3.

Conclusions from Table 3: The main advantage of $\widehat{d}_W$ and $\widetilde{d}_N^{(IR)}$ with respect to $\widehat{d}_{MS}$ and $\widehat{d}_R$ is exhibited in this table: they are robust with respect to smooth trends, especially in the case of long memory processes ($d > 0$). This has already been observed in Bruzaite and Vaiciulis (2008) for IR statistic (and even for certain discontinuous trends). Both those estimators are also robust with respect to seasonal component and this robustness would have been improved if we had chosen m (or scales) as a multiple of the period (which is generally known).
The second good surprise of these simulations is that the adaptive IR estimator $\widetilde{d}_N^{(IR)}$ is also consistent for non Gaussian distributions even if the function Λ in (2.6) and therefore all our results are typically obtained for Gaussian distributions. The case of finite-variance processes is not surprising (see Remark 2). But this is more surprising for infinite variance processes. A first explanation of this was given in Surgailis *et al.* (2007) in the case of i.i.d.r.v. in the domain of attraction of a stable law with index $0 < \alpha < 2$ and skewness parameter $-1 \leq \beta \leq 1$: they concluded that $IR_N(m)$ converges to almost the same limit. The extension to α-stable linear processes of this first explanation should require technical developments but the expression of the IR

statistic (which is bounded in $[0,1]$ for any processes) could allow to apply it to infinite variance processes. Note that the other semiparametric estimators are also consistent in such frame with faster convergence rates notably for the local Whittle estimator.

5 Proofs

Proof of Property 2.1. We proceed in two steps.

Step 1: First, we compute the limit of $\frac{N}{m}\operatorname{Cov}\big(IR_N(jm), IR_N(j'm)\big)$ when N, m and $N/m \to \infty$. As in Surgailis *et al* (2007), define also for all $j=1,\cdots,p$ and $k=1,\cdots,N-3m_j$ (with $m_j = jm$):

$$
\begin{aligned}
Y_{m_j}(k) &:= \frac{1}{V_{m_j}} \sum_{t=k+1}^{k+m_j} (X_{t+m_j} - X_t) \ \ , \text{ with } \ V_{m_j}^2 := \mathrm{E}\Big[\Big(\sum_{t=k+1}^{k+m_j} (X_{t+m_j} - X_t)\Big)^2\Big] \qquad (5.1)\\
\text{and} \quad \eta_{m_j}(k) &:= \frac{|Y_{m_j}(k) + Y_{m_j}(k+m_j)|}{|Y_{m_j}(k)| + |Y_{m_j}(k+m_j)|}. \qquad (5.2)
\end{aligned}
$$

Note that $Y_{m_j}(k) \sim \mathcal{N}(0,1)$ for any k and j and

$$
IR_N(m_j) = \frac{1}{N-3m_j} \sum_{k=0}^{N-3m_j-1} \eta_{m_j}(k) \quad \text{for all } j = 1,\cdots p.
$$

$$
\begin{aligned}
\operatorname{Cov}(IR_N(m_j), IR_N(m_{j'})) &= \frac{1}{N-3m_j}\,\frac{1}{N-3m_{j'}} \sum_{k=0}^{N-3m_j-1} \sum_{k'=0}^{N-3m_{j'}-1} \operatorname{Cov}(\eta_{m_j}(k), \eta_{m_{j'}}(k'))) \\
&= \frac{1}{(\frac{N}{m_j}-3)(\frac{N}{m_{j'}}-3)} \int_{\tau=0}^{\frac{N-1}{m_j}-3} \int_{\tau'=0}^{\frac{N-1}{m_{j'}}-3} \operatorname{Cov}(\eta_{m_j}([m_j\tau]), \eta_{m_{j'}}([m_{j'}\tau']))) \, d\tau\, d\tau'.
\end{aligned}
$$

Now according to (5.20) of the same article, with $\longrightarrow_{FDD}$ denoting the finite distribution convergence when $m \to \infty$,

$$
Y_m([m\tau]) \ \longrightarrow_{FDD} \ Z_d(\tau)
$$

where Z_d is defined in (2.3). Now

$$
\begin{aligned}
Y_{jm}(k) &= \frac{1}{V_{m_j}} \sum_{t=1}^{jm} X_{t+jm+1} - \sum_{t=1}^{jm} X_{t+1} X_t) \\
&= \frac{1}{V_{m_j}} \sum_{i=-(j-1)}^{j-1} (j-|i|) V_m\, Y_m(t+(j+i-1)m).
\end{aligned}
$$

But $V_m^2 \sim c_0 V(d) m^{2d+1}$ when $m \to \infty$ (see (2.20) in Surgailis *et al*, 2007). Therefore we obtain $Y_{jm}([mj\tau]) \sim \frac{1}{j^{d+1/2}} \sum_{i=-(j-1)}^{j-1} (j-|i|) Y_m([mj\tau] + (j+i-1)m)$ when $m \to \infty$ (in distribution) and more generally,

$$
\begin{aligned}
&\big(Y_{jm}([mj\tau]), Y_{j'm}([mj'\tau']) \ \longrightarrow_{FDD} \\
&\quad \Big(\frac{1}{j^{d+1/2}} \sum_{i=-(j-1)}^{j-1} (j-|i|) Z_d(j\tau + j + i - 1)\,,\, \frac{1}{(j')^{d+1/2}} \sum_{i'=-(j'-1)}^{j'-1} (j'-|i'|) Z_d(j'\tau' + j' + i' - 1)\Big), \qquad (5.3)
\end{aligned}
$$

when $m \to \infty$. Hence, obvious computations lead to define for $t \in \mathbb{R}$

$$
\begin{aligned}
Z_d^{(j)}(t) &:= \sum_{i=-(j-1)}^{j-1} (j-|i|) Z_d(t+j+i-1) = \frac{B_{d+0.5}(t+2j) - 2B_{d+0.5}(t+j) + B_{d+0.5}(t)}{\sqrt{|4^{d+0.5}-4|}} \qquad (5.4)\\
\gamma_d^{(j,j')}(t) &:= \operatorname{Cov}\big(\psi(Z_d^{(j)}(0), Z_d^{(j)}(j)), \psi(Z_d^{(j')}(t), Z_d^{(j')}(t+j'))\big). \qquad (5.5)
\end{aligned}
$$

Now, as the function $\psi(x,y) = \frac{|x+y|}{|x|+|y|}$ is a continuous (on $\mathbb{R}^2 \setminus \{0,0\}$) and bounded function (with $0 \leq \psi(x,y) \leq 1$) and since $\eta_{m_j}([mj\tau]) = \psi(Y_{m_j}([m_j\tau]), Y_{m_j}([m_j(\tau+1)]))$, then from (5.3),

$$\begin{aligned}\mathrm{Cov}\big(\eta_{m_j}([m_j\tau]), \eta_{m_{j'}}([m_{j'}\tau'])\big) \quad & \underset{m\to\infty}{\longrightarrow} \quad \mathrm{Cov}\big(\psi(Z_d^{(j)}(j\tau), Z_d^{(j)}(j(\tau+1))), \psi(Z_d^{(j')}(j'\tau'), Z_d^{(j')}(j'(\tau'+1)))\big) \\ & \underset{m\to\infty}{\longrightarrow} \quad \gamma_d^{(j,j')}(j'\tau' - j\tau),\end{aligned}$$

using the stationarity of the process Z_d and therefore of processes $Z_d^{(j)}$ and $Z_d^{(j')}$. Hence, when N, m and $N/m \to \infty$,

$$\begin{aligned}\frac{N}{m}\,\mathrm{Cov}(IR_N(jm), IR_N(j'm)) \quad & \sim \quad \frac{N}{m(\frac{N}{jm}-3)(\frac{N}{j'm}-3)} \\ & \quad \times \int_0^{\frac{N-1}{jm}-3}\int_0^{\frac{N-1}{j'm}-3} \mathrm{Cov}\big(\psi(Z_d^{(j)}(j\,\tau), Z_d^{(j)}(j\,\tau+j)), \psi(Z_d^{(j')}(j'\,\tau'), Z_d^{(j')}(j'\,\tau'+j'))\big)d\tau d\tau' \\ & \sim \quad \frac{mN}{(N-3jm)(N-3j'm)}\int_0^{\frac{N-1}{m}-3j}\int_0^{\frac{N-1}{m}-3j'} \gamma_d^{(j,j')}(s'-s)\,ds\,ds' \\ & \sim \quad \frac{m}{N}\int_{-\frac{N}{m}}^{\frac{N}{m}}\big(\frac{N}{m}-|u|\big)\,\gamma_d^{(j,j')}(u)\,du \\ & \longrightarrow \quad \int_{-\infty}^{\infty}\gamma_d^{(j,j')}(u)\,du =: \sigma_{j,j'}(d).\end{aligned} \tag{5.6}$$

This last limit is obtained, *mutatis mutandis*, from the relation (5.23) Surgailis *et al* (2007), and thus $\gamma_d^{(j,j')}(u) = C\,(u^{-2}\wedge 1)$, implying $\frac{m}{N}\int_{-\frac{N}{m}}^{\frac{N}{m}}|u|\,\gamma_d^{(j,j')}(u)\,du \underset{N,\ m,\ \frac{N}{m}\to\infty}{\longrightarrow} 0$. It achieves the first step of the proof.

Step 2: It remains to prove the multidimensional central limit theorem. Then consider a linear combination of $(IR_N(m_j))_{1\leq j\leq p}$, *i.e.* $\sum_{j=1}^p u_j\,IR_N(m_j)$ with $(u_1,\cdots,u_p)\in\mathbb{R}^p$. For ease of notation, we will restrict our purpose to $p=2$, with $m_i = r_i m$ where $r_1 \leq r_2$ are fixed positive integers. Then with the previous notations and following the notations and results of Theorem 2.5 of Surgailis *et al.* (2007):

$$\begin{aligned}u_1\,IR_N(r_1m) + u_2\,IR_N(r_2m) = u_1(\mathrm{E}[IR_N(r_1m)] + S_K(r_1m) + \widetilde{S}_K(r_1m)) \\ + u_2(\mathrm{E}[IR_N(r_2m)] + S_K(r_2m) + \widetilde{S}_K(r_2m)).\end{aligned}$$

From (5.31) of Surgailis *et al.* (2007), we have $\widetilde{S}_K(m_1) = o(S_K(m_1))$ and $\widetilde{S}_K(m_2) = o(S_K(m_2))$ when $K\to\infty$ and from an Hermitian decomposition $(N/m)^{1/2}(u_1S_K(m_i)+u_2S_K(m_2)) \to_D \mathcal{N}(0,\gamma_K^2)$ as N, m and $N/m\to\infty$ since the cumulants of $(N/m)^{1/2}(u_1S_K(m_i)+u_2S_K(m_2))$ of order greater or equal to 3 converge to 0 (since this result is proved for each $S_K(m_i)$). Moreover, from the previous computations, $\gamma_K^2 \to (u_1^2\sigma_{r_1,r_1}(d) + 2u_1u_2\sigma_{r_1,r_2}(d) + u_2^2\sigma_{r_2,r_2}(d))$ when $K\to\infty$. Therefore the multidimensional central limit theorem is established. $\square$

Property 5.1. *Let X satisfy Assumption $S(d,\beta)$ with $-0.5 < d < 0.5$ and $\beta > 0$. Then, there exists a constant $K(d,\beta) < 0$ depending only on d and β such as*

$$\begin{aligned}\mathrm{E}\big[IR_N(m)\big] &= \Lambda_0(d) + K(d,\beta)\times m^{-\beta} + O\big(m^{-\beta-\varepsilon} + m^{-2d-1}\log(m)\big) && \text{if } -2d+\beta<1, \\ &= \Lambda_0(d) + K(d,\beta)\times m^{-\beta}\,\log(m) \quad + O\big(m^{-\beta}\big) && \text{if } -2d+\beta=1; \\ &= \Lambda_0(d) + O\big(m^{-2d-1}\big) && \text{if } -2d+\beta>1.\end{aligned}$$

Proof of Property 5.1. As in Surgailis *et al* (2007), we can write:

$$\mathrm{E}\big[IR_N(m)\big] = \mathrm{E}\big(\frac{|Y^0+Y^1|}{|Y^0|+|Y^1|}\big) = \Lambda\big(\frac{R_m}{V_m^2}\big) \quad \text{with} \quad \frac{R_m}{V_m^2} := 1 - 2\,\frac{\int_0^\pi f(x)\frac{\sin^6(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx}{\int_0^\pi f(x)\frac{\sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx}.$$

Therefore an expansion of R_m/V_m^2 will provide an expansion of $\mathrm{E}\big[IR_N(m)\big]$ when $m \to \infty$ and the multidimensional CLT (2.8) will be deduced from Delta-method.

Step 1 Let f satisfy Assumption $S(d,\beta)$. Then we are going to establish that there exist positive real numbers C_1 and C_2 specified in (5.7) and (5.8) and such that:

1. if $-1<-2d<1$ and $-2d+\beta<1$, $\dfrac{R_m}{V_m^2}=\rho(d)\ +\ C_1(-2d,\beta)\ \ m^{-\beta}+O\big(m^{-\beta-\varepsilon}+m^{-2d-1}\log m\big)$
2. if $-1<-2d<1$ and $-2d+\beta=1$, $\dfrac{R_m}{V_m^2}=\rho(d)+C_2(1-\beta,\beta)\ m^{-\beta}\log m+O\big(m^{-\beta}\big)$
3. if $-1<-2d<1$ and $-2d+\beta>1$, $\dfrac{R_m}{V_m^2}=\rho(d)+O\big(m^{-2d-1}\big)$.

Indeed under Assumption $S(d,\beta)$ and with $J_j(a,m)$, $j=4,6$, defined in (5.23) of Lemma 5.1 (see below), it is clear that,

$$\frac{R_m}{V_m^2}=1-2\,\frac{J_6(-2d,m)+\frac{c_1}{c_0}J_6(-2d+\beta,m)+O(J_6(-2d+\beta+\varepsilon))}{J_4(-2d,m)+\frac{c_1}{c_0}J_4(-2d+\beta,m)+O(J_4(-2d+\beta+\varepsilon))},$$

since $\displaystyle\int_0^\pi O(x^{-2d+\beta+\varepsilon})\frac{\sin^j(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx=O(J_j(-2d+\beta+\varepsilon))$ for $j=4,6$. Now we follow the results of Lemma 5.1:

1. Let $-1<-2d+\beta<1$. Then for any $\varepsilon>0$,

$$\begin{aligned}\frac{R_m}{V_m^2}&=1-2\,\frac{C_{61}(-2d)m^{1+2d}+C_{62}(-2d)+\frac{c_1}{c_0}\big(C_{61}(-2d+\beta)m^{1+2d-\beta}+C_{62}(-2d+\beta)\big)+O\big(m^{1+2d-\beta-\varepsilon}+\log m\big)}{C_{41}(-2d)m^{1+2d}+C_{42}(-2d)+\frac{c_1}{c_0}\big(C_{41}(-2d+\beta)m^{1+2d-\beta}+C_{42}(-2d+\beta)\big)+O\big(m^{1+2d-\beta-\varepsilon}+\log m\big)}\\&=1-\frac{2}{C_{41}(-2d)}\Big[C_{61}(-2d)+\frac{c_1}{c_0}C_{61}(-2d+\beta)m^{-\beta}\Big]\Big[1-\frac{c_1}{c_0}\frac{C_{41}(-2d+\beta)}{C_{41}(-2d)}m^{-\beta}\Big]+O\big(m^{-\beta-\varepsilon}+m^{-2d-1}\log m\big)\\&=1-\frac{2C_{61}(-2d)}{C_{41}(-2d)}+2\frac{c_1}{c_0}\Big[\frac{C_{61}(-2d)C_{41}(-2d+\beta)}{C_{41}(-2d)C_{41}(-2d)}-\frac{C_{61}(-2d+\beta)}{C_{41}(-2d)}\Big]m^{-\beta}+O\big(m^{-\beta-\varepsilon}+m^{-2d-1}\log m\big).\end{aligned}$$

As a consequence, with $\rho(d)$ defined in (2.7) and C_{j1} defined in Lemma 5.1,

$$\frac{R_m}{V_m^2}=\rho(d)\ +\ C_1(-2d,\beta)\ \ m^{-\beta}+\ O\Big(m^{-\beta-\varepsilon}+m^{-2d-1}\log m\Big)\quad(m\to\infty),\quad\text{with}$$

$$C_1(-2d,\beta):=2\,\frac{c_1}{c_0}\frac{1}{C_{41}^2(-2d)}\big[C_{61}(-2d)C_{41}(-2d+\beta)-C_{61}(-2d+\beta)C_{41}(-2d)\big],\quad(5.7)$$

and numerical experiments proves that $C_1(-2d,\beta)/c_1$ is negative for any $d\in(-0.5,0.5)$ and $\beta>0$.
2. Let $-2d+\beta=1$.
Again with Lemma 5.1,

$$\begin{aligned}\frac{R_m}{V_m^2}&=1-2\frac{[C_{61}(-2d)m^\beta+C'_{61}\frac{c_1}{c_0}log(m\pi)+C_{62}(-2d)+\frac{c_1}{c_0}C'_{62}+O(1)]}{[C_{41}(-2d)m^\beta+C'_{41}\frac{c_1}{c_0}log(m\pi)+C_{42}(-2d)+\frac{c_1}{c_0}C'_{42}+O(1)]}\\&=1-\frac{2}{C_{41}(a)}\big[C_{61}(-2d)+\big(C'_{61}\frac{c_1}{c_0}\log(m)\big)m^{-\beta}\big]\big[1-\big(\frac{C'_{41}}{C_{41}(a)}\frac{c_1}{c_0}\log(m)\big)m^{-\beta}\big]+O\big(m^{-\beta}\big)\\&=1-\frac{2}{C_{41}(-2d)}\Big[C_{61}(-2d)-\frac{c_1}{c_0}\big(\frac{C_{61}(-2d)C'_{41}}{C_{41}(-2d)}-C'_{61}\big)\log(m)\ m^{-\beta}\Big]+O\big(m^{-\beta}\big).\end{aligned}$$

As a consequence,

$$\frac{R_m}{V_m^2}=\rho(d)\ +\ C_2(-2d,\beta)m^{-\beta}\ \log m+O\big(m^{-\beta}\big)\quad(m\to\infty),\quad\text{with}$$

$$C_2(-2d,\beta):=2\,\frac{c_1}{c_0}\frac{1}{C_{41}^2(-2d)}\Big(C'_{41}C_{61}(-2d)-C'_{61}C_{41}(-2d)\Big),\quad(5.8)$$

and numerical experiments proves that $C_2(-2d,\beta)/c_1$ is negative for any $d\in(-0.5,0.5)$ and $\beta=1-2d$.
3. Let $-2d+\beta>1$.
Once again with Lemma 5.1:

$$\begin{aligned}
\frac{R_m}{V_m^2} &= 1-2\frac{\big[C_{61}(-2d)m^{1+2d}+C_{62}(-2d)+\frac{c_1}{c_0}C''_{61}(-2d+\beta)+\frac{c_1}{c_0}C''_{62}(-2d+\beta)m^{1+2d-\beta}+O(1)\big]}{C_{41}(-2d)m^{1+2d}\big[1+\frac{C_{42}(-2d)}{C_{41}(-2d)}m^{-2d-1}+\frac{c_1}{c_0}\frac{C''_{41}(-2d+\beta)}{C_{41}(-2d)}m^{-2d-1}+\frac{c_1}{c_0}\frac{C''_{42}(-2d+\beta)}{C_{41}(-2d)}m^{-\beta}+O(m^{-2d-1})\big]}\\
&= 1-\frac{2}{C_{41}(-2d)}\big[C_{61}(-2d)+O\big(m^{-2d-1}\big)\big]\big[1-O\big(m^{-2d-1}\big)\big]\\
&= 1-\frac{2C_{61}(-2d)}{C_{41}(-2d)}+O\big(m^{-2d-1}\big).
\end{aligned}$$

Note that it is not possible to specify the second order term of this expansion as in both the previous cases. As a consequence,

$$\frac{R_m}{V_m^2}=\rho(d)\ +\ O\big(m^{-2d-1}\big)\quad(m\to\infty). \tag{5.9}$$

Step 2: A Taylor expansion of $\Lambda(\cdot)$ around $\rho(d)$ provides:

$$\Lambda\Big(\frac{R_m}{V_m^2}\Big)\simeq\Lambda\big(\rho(d)\big)+\Big[\frac{\partial\Lambda}{\partial\rho}\Big](\rho(d))\Big(\frac{R_m}{V_m^2}-\rho(d)\Big)+\frac{1}{2}\Big[\frac{\partial^2\Lambda}{\partial\rho^2}\Big](\rho(d))\Big(\frac{R_m}{V_m^2}-\rho(d)\Big)^2. \tag{5.10}$$

Note that numerical experiments show that $\Big[\dfrac{\partial\Lambda}{\partial\rho}\Big](\rho)>0.2$ for any $\rho\in(-1,1)$. As a consequence, using the previous expansions of R_m/V_m^2 obtained in Step 1 and since $\mathrm{E}\big[IR_N(m)\big]=\Lambda\big(R_m/V_m^2\big)$, then

$$\mathrm{E}\big[IR_N(m)\big]=\Lambda_0(d)+\left\{\begin{array}{ll} c_1\,C'_1(d,\beta)\,m^{-\beta}+O\big(m^{-\beta-\varepsilon}+m^{-2d-1}\log m+m^{-2\beta}\big) & \text{if } \beta<1+2d\\ c_1\,C'_2(\beta)m^{-\beta}\log m+O(m^{-\beta}) & \text{if } \beta=1+2d\\ O\big(m^{-2d-1}\big) & \text{if } \beta>1+2d\end{array}\right. ,$$

with $C'_1(d,\beta)<0$ for all $d\in(-0.5,0.5)$ and $\beta\in(0,1+2d)$ and $C'_2(\beta)<0$ for all $0<\beta<2$.

□

Proof of Theorem 1. Using Property 5.1, if $m\simeq C\,N^\alpha$ with $C>0$ and $(1+2\beta)^{-1}\wedge(4d+3)^{-1}<\alpha<1$ then $\sqrt{N/m}\,\big(\mathrm{E}\big[IR_N(m)\big]-\Lambda_0(d)\big)\underset{N\to\infty}{\longrightarrow}0$ and it implies that the multidimensional CLT (2.4) can be replaced by

$$\sqrt{\frac{N}{m}}\Big(IR_N(m_j)-\Lambda_0(d)\Big)_{1\le j\le p}\ \xrightarrow[N\to\infty]{\mathcal{L}}\ \mathcal{N}(0,\Gamma_p(d)). \tag{5.11}$$

It remains to apply the Delta-method with the function Λ_0^{-1} to CLT (5.11). This is possible since the function $d\to\Lambda_0(d)$ is an increasing function such that $\Lambda'_0(d)>0$ and $\big(\Lambda_0^{-1}\big)'(\Lambda_0(d))=1/\Lambda'_0(d)>0$ for all $d\in(-0.5,0.5)$. It achieves the proof of Theorem 1. □

Proof of Proposition 1. For ease of writing we will note $\widehat{\Sigma}_N$ instead of $\widehat{\Sigma}_N(N^\alpha)$ in the sequel. We have $\big(\widetilde{d}_N(m)-\widehat{d}_N(j\,m)\big)_{1\le j\le p}=\widehat{M}_N\big(\widehat{d}_N(j\,m)-d\big)_{1\le j\le p}$ with $\widehat{M}_N$ the orthogonal (for the Euclidian norm $\|\cdot\|_{\widehat{\Sigma}_N}$) projector matrix on $\big((1)_{1\le i\le p}\big)^\perp$ (which is a linear subspace with dimension $p-1$ included in $\mathbb{R}^p$) in $\mathbb{R}^p$, *i.e.* $\widehat{M}_N=J_p(J'_p\widehat{\Sigma}_N^{-1}J_p)^{-1}J'_p\widehat{\Sigma}_N^{-1}$. Now, by denoting $\Sigma_N^{1/2}$ a symmetric matrix such as $\Sigma_N^{1/2}\Sigma_N^{1/2}=\Sigma_N$,

$$\begin{aligned}
\|\big(\widetilde{d}_N(m)-\widehat{d}_N(j\,m)\big)_{1\le j\le p}\|^2_{\widehat{\Sigma}_N} &= \big(\widehat{d}_N(j\,m)-d\big)'_{1\le j\le p}\widehat{M}_N\widehat{\Sigma}_N^{-1}\widehat{M}_N\big(\widehat{d}_N(j\,m)-d\big)_{1\le j\le p}\\
&= Z'_N\widehat{\Sigma}_N^{1/2}\widehat{M}_N\widehat{\Sigma}_N^{-1}\widehat{M}_N\widehat{\Sigma}_N^{1/2}Z_N\\
&= \big(\widehat{A}_NZ_N\big)'\big(\widehat{A}_NZ_N\big)
\end{aligned}$$

with $\widehat{A}_N=\Sigma_N^{-1/2}\widehat{M}_N\widehat{\Sigma}_N^{1/2}$ and Z_N a random vector such as $\sqrt{N/m}Z_N\ \xrightarrow[N\to\infty]{\mathcal{L}}\ \mathcal{N}_p(0,I_p)$ from Theorem 1. But we also have $\widehat{A}_N=\Sigma_N^{-1/2}J_p(J'_p\widehat{\Sigma}_N^{-1}J_p)^{-1}J'_p\widehat{\Sigma}_N^{-1/2}=\widehat{H}_N(\widehat{H}'_N\widehat{H}_N)^{-1}\widehat{H}'_N$ with $\widehat{H}_N=\Sigma_N^{-1/2}J_p$ a matrix of size $(p \times (p-1))$ with rank $p-1$ (since the rank of J_p is $(p-1)$). Hence $\widehat{A}_N$ is an orthogonal projector to the linear subspace of dimension $p-1$ generated by the matrix $\widehat{H}_N$. Now using Cochran Theorem (see for instance Anderson and Styan, 1982), $\sqrt{N/m}\,\widehat{A}_N Z_N$ is asymptotically a Gaussian vector such as $N/m\big(\widehat{A}_N Z_N\big)'\big(\widehat{A}_N Z_N\big) \xrightarrow[N\to\infty]{\mathcal{L}} \chi^2(p-1)$. $\square$

In Property 5.1, a second order expansion of $\mathrm{E}[IR_N(m)]$ can not be specified in the case $\beta > 2d+1$. In the following Property 5.2, we show some inequalities satisfied by $\mathrm{E}[IR_N(m)]$ which will be useful for obtaining the consistency of the adaptive estimator in this case.

Property 5.2. *Let X satisfy Assumption $S(d,\beta)$ with $-0.5 < d < 0.5$, $\beta > 1+2d$. Moreover, suppose that the spectral density of X satisfies Condition (5.13) or (5.14). Then there exists a constant $L > 0$ depending only on $c_0, c_1, c_2, d, \beta, \varepsilon$ such that*

$$\big|\mathrm{E}\big[IR_N(m)\big] - \Lambda_0(d)\big| \geq L\, m^{-2d-1}. \tag{5.12}$$

Proof of Property 5.2. Using the expansion of $J_j(a,m)$, $j=4,6$, for $a>1$ (see Lemma 5.1) and the same computations than in Property 5.1, we obtain:

$$\begin{aligned}
&-\frac{2}{C_{41}^2(-2d)}\Big[\big(C_{62}(-2d)C_{41}(-2d)-C_{42}(-2d)C_{61}(-2d)\big)+\frac{c_1}{c_0}\big(C''_{61}(-2d+\beta)C_{41}(-2d)-C''_{41}(-2d+\beta)C_{61}(-2d)\big)\\
&\qquad+\frac{|c_2|}{c_0}\big(C''_{61}(-2d+\beta+\varepsilon)C_{41}(-2d)+C''_{41}(-2d+\beta+\varepsilon)C_{61}(-2d)\big)\Big]\, m^{-2d-1}(1+o(1))\\
&\qquad\qquad\leq \frac{R_m}{V_m^2}-\rho(d)\leq\\
&-\frac{2}{C_{41}^2(-2d)}\Big[\big(C_{62}(-2d)C_{41}(-2d)-C_{42}(-2d)C_{61}(-2d)\big)+\frac{c_1}{c_0}\big(C''_{61}(-2d+\beta)C_{41}(-2d)-C''_{41}(-2d+\beta)C_{61}(-2d)\big)\\
&\qquad-\frac{|c_2|}{c_0}\big(C''_{61}(-2d+\beta+\varepsilon)C_{41}(-2d)+C''_{41}(-2d+\beta+\varepsilon)C_{61}(-2d)\big)\Big]\, m^{-2d-1}(1+o(1)).
\end{aligned}$$

Now, denote

$$\begin{aligned}
D_0(d) &:= C_{62}(-2d)C_{41}(-2d)-C_{42}(-2d)C_{61}(-2d)=\frac{C_{42}(-2d)C_{41}(-2d)}{48(1-2^{-1+2d})}\big(2^{4+2d}-5-3^{2+2d}\big),\\
D_1(d,\beta) &:= C_{62}(-2d+\beta)C_{41}(-2d)-C_{42}(-2d+\beta)C_{61}(-2d)=\frac{C_{42}(-2d+\beta)C_{41}(-2d)}{128(1-2^{-1+2d})}\big(2^{4+2d}-5-3^{2+2d}\big),\\
D_2(d,\beta,\varepsilon) &:= C''_{61}(-2d+\beta+\varepsilon)C_{41}(-2d)+C''_{41}(-2d+\beta+\varepsilon)C_{61}(-2d).
\end{aligned}$$

Since $-0.5<d<0.5$, $2^{4+2d}-5-3^{2+2d}>0$ and $1-2^{-1+2d}>0$. Moreover, from the sign of the constants presented in Lemma 5.1, we have $D_0(d)\neq 0$ except for $d=0$, $D_1(d,\beta)\neq 0$ except for $d=2\beta$ and $D_2(d,\beta,\varepsilon)>0$ for all $d\in(-0.5,0.5)$, $\beta>0$ and $\varepsilon>0$. Therefore, if $c_0,c_1,c_2,d,\beta,\varepsilon$ are such that

$$K_1 := D_0(d)+\frac{c_1}{c_0}D_1(d,\beta)-\frac{|c_2|}{c_0}D_2(d,\beta,\varepsilon)>0 \tag{5.13}$$

$$\text{or}\quad K_2 := D_0(d)+\frac{c_1}{c_0}D_1(d,\beta)+\frac{|c_2|}{c_0}D_2(d,\beta,\varepsilon)<0. \tag{5.14}$$

and from the signs of $D_0(d)$, $D_1(d,\beta)$ and $D_2(d,\beta,\varepsilon)$, when (d,β,ε) is fixed, these conditions are not impossible but hold following the values of $\frac{c_1}{c_0}$ and $\frac{|c_2|}{c_0}$. Then $\frac{R_m}{V_m^2}-\rho(d)\leq -\frac{K_1}{C_{41}^2(-2d)}\,m^{-2d-1}$ or $\frac{R_m}{V_m^2}-\rho(d)\geq -\frac{K_2}{C_{41}^2(-2d)}\,m^{-2d-1}$ for m large enough following (5.13) or (5.14) holds. Then, if (5.13) holds, since $\mathrm{E}[IR_N(m)]=\Lambda(\frac{R_m}{V_m^2})$, since the function $r\to\Lambda(r)$ is an increasing and $\mathcal{C}^1$ function and since $\mathrm{E}[IR_N(m)]=\Lambda\big(\frac{R_m}{V_m^2}\big)$ then when m large enough, from a Taylor expansion,

$$\mathrm{E}[IR_N(m)]\leq\Lambda\Big(\rho(d)-\frac{K_1}{C_{41}^2(-2d)}\,m^{-2d-1}\Big)\quad\Longrightarrow\quad \mathrm{E}[IR_N(m)]\leq\Lambda_0(d)-\frac{1}{2}\Lambda'(\rho(d))\frac{K_1}{C_{41}^2(-2d)}\,m^{-2d-1}.$$

Now following the same process if (5.14) holds, we deduce inequality (5.12). $\square$

Proof of Proposition 2. Let $\varepsilon > 0$ be a fixed positive real number, such that $\alpha^* + \varepsilon < 1$.

I. First, a bound of $\Pr(\widehat{\alpha}_N \leq \alpha^* + \varepsilon)$ is provided. Indeed,

$$\begin{aligned}
\Pr\left(\widehat{\alpha}_N \leq \alpha^* + \varepsilon\right) &\geq \Pr\Big(\widehat{Q}_N(\alpha^* + \varepsilon/2) \leq \min_{\alpha \geq \alpha^* + \varepsilon \text{ and } \alpha \in \mathcal{A}_N} \widehat{Q}_N(\alpha)\Big) \\
&\geq 1 - \Pr\Big(\bigcup_{\alpha \geq \alpha^* + \varepsilon \text{ and } \alpha \in \mathcal{A}_N} \widehat{Q}_N(\alpha^* + \varepsilon/2) > \widehat{Q}_N(\alpha)\Big) \\
&\geq 1 - \sum_{k=[(\alpha^*+\varepsilon)\log N]}^{\log[N/p]} \Pr\Big(\widehat{Q}_N(\alpha^* + \varepsilon/2) > \widehat{Q}_N\big(\frac{k}{\log N}\big)\Big). \qquad (5.15)
\end{aligned}$$

But, for $\alpha \geq \alpha^* + \varepsilon$,

$$\begin{aligned}
&\Pr\Big(\widehat{Q}_N(\alpha^* + \varepsilon/2) > \widehat{Q}_N(\alpha)\Big) \\
&\quad = \Pr\Big(\big\|\big(\widehat{d}_N(i\, N^{\alpha^*+\varepsilon/2})\big)_{1\leq i\leq p} - \widetilde{d}_N(N^{\alpha^*+\varepsilon/2})\big\|^2_{\widehat{\Sigma}_N(N^{\alpha^*+\varepsilon/2})} > \big\|\big(\widehat{d}_N(i\, N^{\alpha}) - \widetilde{d}_N(N^{\alpha})\big)_{1\leq i\leq p}\big\|^2_{\widehat{\Sigma}_N(N^{\alpha})}\Big)
\end{aligned}$$

with $\|X\|^2_\Omega = X'\,\Omega^{-1}\,X$. Set $Z_N(\alpha) = \frac{N}{N^\alpha}\,\big\|\big(\widehat{d}_N(i\, N^{\alpha})\big)_{1\leq i\leq p} - \widetilde{d}_N(N^{\alpha})\big\|^2_{\widehat{\Sigma}_N(N^{\alpha})}$. Then,

$$\begin{aligned}
\Pr\Big(\widehat{Q}_N(\alpha^* + \varepsilon/2) > \widehat{Q}_N(\alpha)\Big) &= \Pr\Big(Z_N(\alpha^* + \varepsilon/2) > N^{\alpha-(\alpha^*+\varepsilon/2)}\, Z_N(\alpha)\Big) \\
&\leq \Pr\Big(Z_N(\alpha^* + \varepsilon/2) > N^{(\alpha-(\alpha^*+\varepsilon/2))/2}\Big) + \Pr\Big(Z_N(\alpha) < N^{-(\alpha-(\alpha^*+\varepsilon/2))/2}\Big).
\end{aligned}$$

From Proposition 1, for all $\alpha > \alpha^*$, $Z_N(\alpha) \xrightarrow[N\to\infty]{\mathcal{L}} \chi^2(p-1)$. As a consequence, for N large enough,

$$\Pr\Big(Z_N(\alpha) \leq N^{-(\alpha-(\alpha^*+\varepsilon/2))/2}\Big) \leq \frac{2}{2^{(p-1)/2}\Gamma((p-1)/2)} \cdot N^{-(\frac{p-1}{2})\frac{(\alpha-(\alpha^*+\varepsilon/2))}{2}}.$$

Moreover, from Markov inequality and with N large enough,

$$\begin{aligned}
\Pr\Big(Z_N(\alpha^* + \varepsilon/2) > N^{(\alpha-(\alpha^*+\varepsilon/2))/2}\Big) &\leq 2\,\Pr\Big(\exp(\sqrt{\chi^2(p-1)}) > \exp\big(N^{(\alpha-(\alpha^*+\varepsilon/2))/4}\big)\Big) \\
&\leq 2\,\mathrm{E}(\exp(\sqrt{\chi^2(p-1)}))\,\exp\big(-N^{(\alpha-(\alpha^*+\varepsilon/2))/4}\big).
\end{aligned}$$

We deduce that there exists $M_1 > 0$ not depending on N, such that for large enough N,

$$\Pr\Big(\widehat{Q}_N(\alpha^* + \varepsilon/2) > \widehat{Q}_N(\alpha)\Big) \leq M_1\,\exp\big(-N^{(\alpha-(\alpha^*+\varepsilon/2))/4}\big).$$

since $\mathrm{E}(\exp(\sqrt{\chi^2(p-1)})) < \infty$ does not depend on N. Thus, the inequality (5.15) becomes, with $M_2 > 0$ and for N large enough,

$$\begin{aligned}
\Pr\left(\widehat{\alpha}_N \leq \alpha^* + \varepsilon\right) &\geq 1 - M_1\, e^{-N^{\varepsilon/8}} \sum_{k=0}^{\log[N/p]-[(\alpha^*+\varepsilon)\log N]} \exp\big(-N^{\frac{k}{4\log N}}\big) \\
&\geq 1 - M_2\, e^{-N^{\varepsilon/8}}. \qquad (5.16)
\end{aligned}$$

II. Secondly, a bound of $\Pr(\widehat{\alpha}_N \geq \alpha^* - \varepsilon)$ can also be computed. Following the previous arguments and notations,

$$\begin{aligned}
\Pr\left(\widehat{\alpha}_N \geq \alpha^* - \varepsilon\right) &\geq \Pr\Big(\widehat{Q}_N(\alpha^* + \frac{1-\alpha^*}{2\alpha^*}\varepsilon) \leq \min_{\alpha \leq \alpha^* - \varepsilon \text{ and } \alpha \in \mathcal{A}_N} \widehat{Q}_N(\alpha)\Big) \\
&\geq 1 - \sum_{k=2}^{[(\alpha^*-\varepsilon)\log N]+1} \Pr\Big(\widehat{Q}_N(\alpha^* + \frac{1-\alpha^*}{2\alpha^*}\varepsilon) > \widehat{Q}_N\big(\frac{k}{\log N}\big)\Big), \qquad (5.17)
\end{aligned}$$

and as above, with $Z_N(\alpha)=\frac{N}{N^\alpha}\left\|\big(\widehat{d}_N(i\,N^\alpha)-\widetilde{d}_N(N^\alpha)\big)_{1\le i\le p}\right\|^2_{\widehat{\Sigma}_N(N^\alpha)}$,

$$\Pr\Big(\widehat{Q}_N(\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon)>\widehat{Q}_N(\alpha)\Big)=\Pr\Big(Z_N(\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon)>N^{\alpha-(\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon)}Z_N(\alpha)\Big). \tag{5.18}$$

• if $\beta\le 2d+1$, with $\alpha<\alpha^*=(1+2\beta)^{-1}$, from Property 5.1 and with $C\neq 0$, for $1\le i\le p$,

$$\sqrt{\frac{N}{N^\alpha}}\big(\mathrm{E}\big[IR(i\,N^\alpha)\big]-\Lambda_0(d)\big)\simeq C\,i^{-(1-\alpha^*)/2\alpha^*}\,N^{(\alpha^*-\alpha)/2\alpha^*}(\log N)^{\mathbf{1}_{\beta=2d+1}}$$
$$\Longrightarrow\sqrt{\frac{N}{N^\alpha}}\big(\Lambda_0^{-1}(\mathrm{E}\big[IR(i\,N^\alpha)\big])-d\big)\simeq C'\,i^{-(1-\alpha^*)/2\alpha^*}\,N^{(\alpha^*-\alpha)/2\alpha^*}(\log N)^{\mathbf{1}_{\beta=2d+1}} \tag{5.19}$$

with $C'\neq 0$, since $\Lambda_0(d)>0$ for all $d\in(-0.5,0.5)$. We deduce:

$$\sqrt{\frac{N}{N^\alpha}}\Big(\widehat{d}_N(i\,N^\alpha)-d\Big)_{1\le i\le p}\simeq C''\,N^{(\alpha^*-\alpha)/2\alpha^*}(\log N)^{\mathbf{1}_{\beta=2d+1}}\big(i^{-(1-\alpha^*)/2\alpha^*}\big)_{1\le i\le p}+\big(\widehat{\varepsilon}_N(i\,N^\alpha)\big)_{1\le i\le p},$$

with $C''\neq 0$ and $\big(\widehat{\varepsilon}_N(i\,N^\alpha)\big)_{1\le i\le p}\xrightarrow[N\to\infty]{\mathcal{L}}\mathcal{N}\Big(0,(\Lambda_0'(d))^{-2}\,\Gamma_p(d)\Big)$ from Proposition 1. Now from the definition of $\widetilde{d}_N(N^\alpha)$, we have $\big(\widehat{d}_N(i\,N^\alpha)-\widetilde{d}_N(N^\alpha)\big)_{1\le i\le p}=\widehat{M}_N\big(\widehat{d}_N(i\,N^\alpha)-d\big)_{1\le i\le p}$ with $\widehat{M}_N$ the orthogonal projector matrix on $(1)^{\perp}_{1\le i\le p}$.

As a consequence, for $\alpha<\alpha^*-\varepsilon$ and with the inequality $\|a-b\|^2\ge\frac12\|a\|^2-\|b\|^2$,

$$Z_N(\alpha)\quad\ge\quad\frac12(C'')^2\,N^{\frac{\alpha^*-\alpha}{\alpha^*}}(\log^2 N)^{\mathbf{1}_{\beta=2d+1}}\Big\|\widehat{M}_N\big(i^{-\frac{1-\alpha^*}{2\alpha^*}}\big)_{1\le i\le p}\Big\|_{\widehat{\Sigma}_N(N^\alpha)}-\|\widehat{M}_N\widehat{\varepsilon}_N(i\,N^\alpha))\|^2_{\widehat{\Sigma}_N(N^\alpha)}.$$

Now, it is clear that $\|\widehat{M}_N\widehat{\varepsilon}_N(i\,N^\alpha))\|^2_{\widehat{\Sigma}_N(N^\alpha)}\le\|\widehat{\varepsilon}_N(i\,N^\alpha))\|^2_{\widehat{\Sigma}_N(N^\alpha)}\le C_1$ when N large enough, with $C_1>0$ not depending on N. Moreover the vector $\big(i^{-\frac{1-\alpha^*}{2\alpha^*}}\big)_{1\le i\le p}$ is not in the subspace $(1)_{1\le i\le p}$ and therefore $\Big\|\widehat{M}_N\big(i^{-\frac{1-\alpha^*}{2\alpha^*}}\big)_{1\le i\le p}\Big\|_{\widehat{\Sigma}_N(N^\alpha)}\ge C_2$ for N large enough with $C_2>0$. We deduce that there exists $D>0$ such that for N large enough and $\alpha<\alpha^*-\varepsilon$,

$$Z_N(\alpha)\quad\ge\quad D\,N^{\frac{\alpha^*-\alpha}{\alpha^*}}(\log^2 N)^{\mathbf{1}_{\beta=2d+1}}.$$

Therefore, since $N^{\frac{\alpha^*-\alpha}{\alpha^*}}\xrightarrow[N\to\infty]{}\infty$ when $\alpha<\alpha^*-\varepsilon$,

$$\Pr\Big(Z_N(\alpha)\ge\frac12 D\,N^{\frac{\alpha^*-\alpha}{\alpha^*}}\Big)\xrightarrow[N\to\infty]{}1.$$

Then, the relation (5.18) becomes for $\alpha<\alpha^*-\varepsilon$ and N large enough,

$$\begin{aligned}\Pr\Big(\widehat{Q}_N(\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon)>\widehat{Q}_N(\alpha)\Big)\quad&\le\quad\Pr\Big(\chi^2(p-1)\ge\big(\frac12\,D\,N^{\frac{\alpha^*-\alpha}{\alpha^*}}\big)\,N^{\alpha-(\alpha^*+\frac{1-\alpha^*}{2\alpha^*}\varepsilon)}\Big)\\&\le\quad\Pr\Big(\chi^2(p-1)\ge\frac{D}{2}\,N^{\frac{1-\alpha^*}{2\alpha^*}(2(\alpha^*-\alpha)-\varepsilon)}\Big)\\&\le\quad M_2\,N^{-(\frac{p-1}{2})\frac{1-\alpha^*}{2\alpha^*}\,\varepsilon},\end{aligned}$$

with $M_2>0$, because $\frac{1-\alpha^*}{2\alpha^*}(2(\alpha^*-\alpha)-\varepsilon)\ge\frac{1-\alpha^*}{2\alpha^*}\varepsilon$ for all $\alpha\le\alpha^*-\varepsilon$. Hence, from the inequality (5.17), for large enough N,

$$\Pr\big(\widehat{\alpha}_N\ge\alpha^*-\varepsilon\big)\ge 1-M_2\,\log N\,N^{-(p-1)\frac{1-\alpha^*}{4\alpha^*}\,\varepsilon}. \tag{5.20}$$

• if $\beta>2d+1$, with $\alpha<\alpha^*=(4d+3)^{-1}$ and from Property 5.2, we obtain an inequality instead of (5.19):

$$\Big|\Lambda_0^{-1}\big(\mathrm{E}[IR_N(m)]\big)-d\Big|\ge\frac12(\Lambda_0(d))^{-1}L\,m^{-2d-1}$$

since the function $x\mapsto\Lambda_0^{-1}(x)$ is an increasing an $\mathcal{C}^1$ function, using a Taylor expansion. Therefore for $1\le i\le p$,

$$\sqrt{\frac{N}{N^\alpha}}\Big|\Lambda_0^{-1}\big(\mathrm{E}\big[IR(i\,N^\alpha)\big]\big)-d\Big|\ge\frac12(\Lambda_0(d))^{-1}L\,i^{-(1-\alpha^*)/2\alpha^*}\,N^{(\alpha^*-\alpha)/2\alpha^*}. \tag{5.21}$$

Now, as previously and with the same notation,

$$\big(\widehat{d}_N(i\,N^\alpha)\;-\;\widetilde{d}_N(N^\alpha)\big)_{1\le i\le p}\;\;\simeq\;\;\widehat{M}_n\big(\Lambda_0^{-1}\Big(\mathrm{E}\big[IR(i\,N^\alpha)\big]\big)\;-\;d\Big)_{1\le i\le p}\;+\;\widehat{M}_n\big(\widehat{\varepsilon}_N(i\,N^\alpha)\big)_{1\le i\le p}. \quad (5.22)$$

Now plugging (5.21) in (5.22) and following the same steps of the proof in the case $\beta \le 2d+1$, the same kind of bound (5.20) can be obtained.

Finally, the inequalities (5.16) and (5.20) imply that $\Pr\big(|\widehat{\alpha}_N - \alpha^*| \ge \varepsilon\big) \underset{N\to\infty}{\longrightarrow} 0.$ $\square$

Proof of Theorem 2. The results of Theorem 2 can be easily deduced from Theorem 1 and Proposition 2 (and its proof) by using conditional probabilities. $\square$

Proof of Proposition 3. Proposition 3 can be deduced from Theorem 2 using the same kind of proof than in Proposition 1 and conditional distributions. $\square$

Lemma 5.1. *For $j=4,6$, denote*

$$J_j(a,m) := \int_0^\pi x^a \frac{\sin^j(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx. \quad (5.23)$$

Then, we have the following expansion when $m\to\infty$:

1. *if $-1<a<1$,* $J_j(a,m) = C_{j1}(a)\,m^{1-a} + C_{j2}(a) + O\big(m^{-1-(a\wedge 0)}\big)$;
2. *if $a=1$,* $J_j(a,m) = C'_{j1}\,\log(m) + C'_{j2} + O\big(m^{-1}\big)$;
3. *if $a>1$,* $J_j(a,m) = C''_{j1}(a) + O\big(m^{1-a}+m^{-2}\big)$,

where constants $C_{j1}(a)$, $C_{j2}(a)$, $C'_{j1}(a)$, $C'_{j2}(a)$ and $C''_{j1}(a)$ are specified in the following proof.

Proof of Lemma 5.1. 1. let $-1<a<1$.
We begin with the expansion of $J_4(a,m)$. First, decompose $J_4(a,m)$ as follows

$$J_4(a,m) = 2^{a+1}\int_0^{\frac{\pi}{2}} y^a\sin^4(my)\Big[\frac{1}{\sin^2(y)}-\frac{1}{y^2}\Big]dy + \int_0^\pi \frac{x^a}{(\frac{x}{2})^2}\sin^4(\frac{mx}{2})dx. \quad (5.24)$$

Using integrations by parts and $\sin^4(\frac{x}{2}) = \sin^2(\frac{x}{2}) - \frac{1}{4}\sin^2(x) = \frac{1}{8}\big(3-4\cos(y)+\cos(2y)\big)$, we obtain for $m\to\infty$:

$$\begin{aligned}\int_0^\pi \frac{x^a}{(\frac{x}{2})^2}\sin^4(\frac{mx}{2})dx &= 4\,m^{1-a}\Big((1-\frac{1}{2^{1+a}})\int_0^\infty \frac{\sin^2(\frac{y}{2})}{y^{2(\frac{1-a}{2})+1}}dy - \frac{1}{8}\int_{m\pi}^\infty y^{a-2}\big(3-4\cos(y)+\cos(2y)\big)dy\Big)\\ &= \frac{\pi(1-\frac{1}{2^{1+a}})}{(1-a)\Gamma(1-a)\sin(\frac{(1-a)\pi}{2})}\,m^{1-a} - 3\,\frac{1}{2(1-a)}\pi^{a-1} + O(m^{-1})\end{aligned}$$

where the left right side term of the last relation is obtained by integration by parts and the left side term is deduced from the following relation (see Doukhan *et al.* 2003, p. 31)

$$\int_0^\infty y^{-\alpha}\;\sin(y)\;dy \;=\; \frac{1}{2}\,\frac{\pi}{\Gamma(\alpha)\sin(\pi(\frac{\alpha}{2}))} \quad \text{for } 0<\alpha<2. \quad (5.25)$$

Moreover, with the linearization of $\sin^4 u$ and Taylor expansions $\frac{1}{\sin^2(y)}-\frac{1}{y^2} \underset{y\to 0}{\sim} \frac{1}{3}$ and $\frac{1}{y^3}-\frac{\cos(y)}{\sin^3(y)} \underset{y\to 0}{\sim} \frac{y}{15}$,

$$2^{a+1}\int_0^{\frac{\pi}{2}} y^a\sin^4(my)\Big[\frac{1}{\sin^2(y)}-\frac{1}{y^2}\Big]dy = 3\;\frac{2^{a+1}}{8}\int_0^{\frac{\pi}{2}} y^a\big[\frac{1}{\sin^2(y)}-\frac{1}{y^2}\big]dy + O\big(m^{-1-(a\wedge 0)}\big). \quad (5.26)$$

Finally, by replacing this expansion in (5.24), one deduces

$$J_4(a,m) = \int_0^\pi x^a\frac{\sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx = C_{41}(a)\,m^{1-a} + C_{42}(a) + O\big(m^{-1-(a\wedge 0)}\big) \quad (m\to\infty), \text{with}$$

$$C_{41}(a) := \frac{\pi(1-\frac{1}{2^{1+a}})}{(1-a)\Gamma(1-a)\sin(\frac{(1-a)\pi}{2})} \;\text{ and } C_{42}(a) := \frac{3}{2^{2-a}}\int_0^{\frac{\pi}{2}} y^a\big[\frac{1}{\sin^2(y)}-\frac{1}{y^2}\big]dy - \frac{3}{2(1-a)}\pi^{a-1}. \quad (5.27)$$

Note that $C_{41}(a) > 0$ and $C_{42}(a) < 0$ for all $0 < a < 1$, $C_{42}(a) > 0$ for all $-1 < a < 0$, $C_{42}(0) = 0$.

A similar expansion procedure of $J_6(a, m)$ with $\sin^6(\frac{mx}{2})$ instead of $\sin^4(\frac{mx}{2})$ can be provided. As previously with $\sin^6(\frac{y}{2}) = \frac{1}{32}\big(10 - 15\cos(y) + 6\cos(2y) - \cos(3y)\big)$, when $m \to \infty$,

$$J_6(a,m) = C_{61}(a)\, m^{1-a} + C_{62}(a) + O\big(m^{-1-(a\wedge 0)}\big),$$

$$\text{with } C_{61}(a) := \frac{\pi(15 + 3^{1-a} - 2^{1-a}6)}{16(1-a)\Gamma(1-a)\sin(\frac{\pi}{2}(1-a))} \text{ and } C_{62}(a) := \frac{5}{6}\, C_{42}(a).$$

Moreover it is clear that $C_{61}(a) > 0$.

2. let $a = 1$.
When $m \to \infty$ we obtain the following expansion:

$$\int_0^\pi \frac{x\sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx \;=\; \frac{1}{2}\Big(\int_0^{m\pi} \frac{\sin(2x) - 2x}{2x^2}dx - 4\int_0^{m\pi} \frac{\sin(x) - x}{x^2}dx\Big) + 4\int_0^{\frac{\pi}{2}} y\sin^4(my)\Big(\frac{1}{\sin^2(y)} - \frac{1}{y^2}\Big)dy$$

But,

$$\int_0^{m\pi} \frac{\sin(2x) - 2x}{2x^2}dx - 4\int_0^{m\pi} \frac{\sin(x) - x}{x^2}dx = \frac{3}{2}\Big(\log(m\pi) + \int_1^\infty \frac{\sin y}{y^2}dy + \int_0^1 \frac{\sin y - y}{y^2}dy\Big) + O(m^{-1}).$$

Moreover from previous computations (see the case $a < 1$),

$$\int_0^{\frac{\pi}{2}} y\sin^4(my)\Big(\frac{1}{\sin^2(y)} - \frac{1}{y^2}\Big)dy = \frac{3}{8}\int_0^{\frac{\pi}{2}} y\Big(\frac{1}{\sin^2(y)} - \frac{1}{y^2}\Big)dy + O(m^{-1}).$$

As a consequence, when $m \to \infty$,

$$\int_0^\pi \frac{x\sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx = C'_{41}\,\log(m) + C'_{42} + O\big(m^{-1}\big), \quad \text{with} \quad C'_{41} := \frac{3}{2} \quad \text{and}$$

$$C'_{42} := \frac{3}{2}\Big(\log(\pi) + \int_0^{\frac{\pi}{2}} y\Big(\frac{1}{\sin^2(y)} - \frac{1}{y^2}\Big)dy + \int_1^\infty \frac{\sin y}{y^2}dy + \int_0^1 \frac{\sin y - y}{y^2}dy\Big).$$

Note that $C'_{41} > 0$ and $C'_{42} \simeq 2.34 > 0$.

In the same way , we obtain the following expansions when $m \to \infty$,

$$\int_0^\pi \frac{x\sin^6(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx = C'_{61}\,\log(m) + C'_{62} + O\big(m^{-1}\big) \quad \text{with} \quad C'_{61} := \frac{5}{4} \quad \text{and}$$

$$C'_{62} := \frac{5}{4}\log(\pi) + \frac{5}{4}\int_0^{\frac{\pi}{2}} y\Big(\frac{1}{\sin^2(y)} - \frac{1}{y^2}\Big)dy + \frac{1}{8}\int_1^\infty \frac{1}{y}\Big(-\cos(3y) + 6\cos(2y) - 15\cos(y)\Big)dy + 4\int_0^1 \frac{1}{y}\sin^6(\frac{y}{2})dy.$$

Note again that $C'_{61} > 0$ and numerical experiments show that $C'_{62} > 0$.

3. Let $a > 1$. Then, with the linearization of $\sin^4(u)$,

$$\begin{aligned}\int_0^\pi \frac{x^a\sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx &= \frac{3}{8}\int_0^\pi \frac{x^a}{\sin^2(\frac{x}{2})}dx - \frac{1}{2}\int_0^\pi \frac{x^a}{\sin^2(\frac{x}{2})}\cos(mx)dx + \frac{1}{8}\int_0^\pi \frac{x^a}{\sin^2(\frac{x}{2})}\cos(2mx)dx \\ &= C''_{41}(a) + \frac{1}{m}\int_0^\pi \Big(\frac{\sin(mx)}{2} - \frac{\sin(2mx)}{16}\Big)\Big(g(x) + h(x)\Big)dx, \end{aligned} \qquad (5.28)$$

with: $g(x) = \Big(\frac{ax^{a-1}}{\sin^2(\frac{x}{2})} - 4ax^{a-3}\Big) - \Big(\frac{x^a\cos(\frac{x}{2})}{\sin^3(\frac{x}{2})} - 8x^{a-3}\Big)$ and $h(x) = (4a - 8)x^{a-3}$.

First, if $1 < a$, with an integration by parts,

$$\frac{1}{m}\int_0^\pi \Big(\frac{\sin(mx)}{2} - \frac{\sin(2mx)}{16}\Big)h(x)dx \;=\; O\big(m^{1-a} + m^{-2}\big). \qquad (5.29)$$

Moreover,

$$\frac{1}{m}\int_0^{\pi}\Big(\frac{\sin(mx)}{2}-\frac{\sin(2mx)}{16}\Big)g(x)dx$$
$$=\Big(\frac{1}{32}-\frac{(-1)^m}{2}\Big)\big(a\pi^2-4a+8\big)\pi^{a-3}\frac{1}{m^2}-\frac{1}{m^2}\int_0^{\pi}\Big(-\frac{\cos(mx)}{2}+\frac{\cos(2mx)}{32}\Big)\,g'(x)dx$$

since $g(x)\underset{x=0^+}{\sim}\frac{a}{3}\,x^{a-1}$ and $g'(x)\underset{x=0^+}{\sim}\frac{a(a-1)}{3}x^{a-2}$. Therefore, if $1<a$,

$$\frac{1}{m}\int_0^{\pi}\Big(\frac{\sin(mx)}{2}-\frac{\sin(2mx)}{16}\Big)g(x)dx \;=\; O\big(m^{-2}\big).$$

In conclusion, for $1<a$ we deduce,

$$\int_0^{\pi}\frac{x^a\sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx \;=\; C''_{41}(a)\;+\;O\big(m^{1-a}\;+\;m^{-2}\big)\quad\text{with}\quad C''_{41}(a)\;:=\;\frac{3}{8}\int_0^{\pi}\frac{x^a}{\sin^2(\frac{x}{2})}dx\;>\;0.$$

Similarly, for $1<a$ we deduce,

$$\int_0^{\pi}\frac{x^a\sin^6(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx=C''_{61}(a)+O\big(m^{1-a}+m^{-2}\big)\quad\text{with}\quad C''_{61}(a):=\frac{5}{16}\int_0^{\pi}\frac{x^a}{\sin^2(\frac{x}{2})}dx=\frac{5}{6}\,C''_{41}(a)>0.$$

$\square$

Acknowledgments. The authors are grateful to both the referees for their very careful reading and many relevant suggestions and corrections that strongly improve the content and the form of the paper.

References

[1] Anderson, T.W. and Styan, G.P.H. (1982). Cochran's theorems, rank additivity and tripotent matrices. In *Statistics and probability: essays in honor of C.R. Rao*, 1-23, North-Holland, Amsterdam-New York.

[2] Bardet, J.M. (2000). Testing for the presence of self-similarity of Gaussian time series having stationary increments, *J. Time Ser. Anal.* , 21, 497-516.

[3] Bardet, J.M., Bibi H. and Jouini, A. (2008). Adaptive wavelet-based estimator of the memory parameter for stationary Gaussian processes, *Bernoulli*, 14, 691-724.

[4] Bardet, J.M., Lang, G., Oppenheim, G., Philippe, A., Stoev, S. and Taqqu, M.S. (2003). Semiparametric estimation of the long-range dependence parameter: a survey. In *Theory and applications of long-range dependence*, Birkhäuser Boston, 557-577.

[5] Bardet J.M. and Surgailis, D. (2011). Measuring the roughness of random paths by increment ratios, *Bernoulli*, 17, 749-780.

[6] Bruzaite, K. and Vaiciulis, M. (2008). The increment ratio statistic under deterministic trends. *Lith. Math. J.* 48, 256-269.

[7] Dahlhaus, R. (1989) Efficient parameter estimation for self-similar processes, *Ann. Statist.*, 17, 1749-1766.

[8] Doukhan, P., Oppenheim, G. and Taqqu M.S. (Editors) (2003). *Theory and applications of long-range dependence*, Birkhäuser.

[9] Fox, R. and Taqqu, M.S. (1986). Large-sample properties of parameter estimates for strongly dependent Gaussian time series. *Ann. Statist.* 14, 517-532.

[10] Geweke, J. and Porter-Hudak, S. (1983), The estimation and application of long-memory time-series models, *J. Time Ser. Anal.*, 4, 221-238.

[11] Giraitis, L., Robinson P.M., and Samarov, A. (1997). Rate optimal semiparametric estimation of the memory parameter of the Gaussian time series with long range dependence, *J. Time Ser. Anal.*, 18, 49-61.

[12] Giraitis, L. and Surgailis, D. (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its applications to the asymptotic normality of Whittle estimate. *Prob. Th. and Rel. Field.* 86, 87-104.

[13] Giraitis, L. and Taqqu, M.S. (1999). Whittle estimator for finite-variance non-Gaussian time series with long memory. *Ann. Statist.* 27, 178–203.

[14] Hurst, H. E. (1951) Long-term storage capacity of reservoirs, *Trans, Am. Soc. Civil Eng*, 116, 770-779.

[15] Ho H.C. and Hsing T. (1997). Limit theorems for functionals of moving averages, *Ann. Probab.* 25, 1636–1669.

[16] Moulines, E., Roueff, F. and Taqqu, M.S. (2007). On the spectral density of the wavelet coefficients of long memory time series with application to the log-regression estimation of the memory parameter, *J. Time Ser. Anal.*, 28, 155-187.

[17] Moulines, E. and Soulier, P. (2003). Semiparametric spectral estimation for fractionnal processes, In P. Doukhan, G. Openheim and M.S. Taqqu editors, *Theory and applications of long-range dependence*, 251-301, Birkhäuser, Boston.

[18] Robinson, P.M. (1995a). Log-periodogram regression for time series with long-range dependence, *Ann. Statist.*, 23, 1048-1072.

[19] Robinson, P.M. (1995b). Gaussian semiparametric estimation of long range dependence, *Ann. Statist.*, 23, 1630-1661.

[20] Surgailis, D., Teyssière, G. and Vaičiulis, M. (2007) The increment ratio statistic. *J. Multiv. Anal.* 99, 510-541.

[21] Vaiciulis, M. (2009). An estimator of the tail index based on increment ratio statistics. *Lith. Math. J.* 49, 222-233.

[22] Veitch, D., Abry, P. and Taqqu, M.S. (2003). On the Automatic Selection of the Onset of Scaling, *Fractals* 11, 377-390.

	Model	Estimates	$p=5$	$p=10$	$p=15$	$p=20$
$N=10^3$	fGn ($H=d+1/2$)	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	**0.088***	0.094	0.101	0.111
		mean($\widetilde{m}_N$)	11.8	12.5	16.0	19.4
		$\widehat{proba}$	0.93	0.89	0.86	0.85
	FARIMA$(0,d,0)$	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	0.112	0.099	**0.094***	0.107
		mean($\widetilde{m}_N$)	13.9	12.5	14.6	17.9
		$\widehat{proba}$	0.94	0.92	0.88	0.86
	FARIMA$(1,d,1)$	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	0.141	**0.136***	0.140	0.149
		mean($\widetilde{m}_N$)	15.2	15.0	18.2	21.1
		$\widehat{proba}$	0.94	0.89	0.86	0.82
	$X^{(d,\beta)}$, $\beta=1$	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	0.122	**0.112***	0.121	0.123
		mean($\widetilde{m}_N$)	14.1	13.8	16.2	20.0
		$\widehat{proba}$	0.91	0.90	0.87	0.85

	Model	Estimates	$p=5$	$p=10$	$p=15$	$p=20$
$N=10^4$	fGn ($H=d+1/2$)	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	0.030	0.022	0.019	**0.018***
		mean($\widetilde{m}_N$)	13.7	10.3	9.4	8.9
		$\widehat{proba}$	0.95	0.89	0.87	0.84
	FARIMA$(0,d,0)$	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	0.039	0.034	0.033	**0.031***
		mean($\widetilde{m}_N$)	11.5	9.0	8.0	7.2
		$\widehat{proba}$	0.95	0.90	0.88	0.82
	FARIMA$(1,d,1)$	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	0.067	0.062	**0.061***	**0.061***
		mean($\widetilde{m}_N$)	18.1	15.9	13.8	13.3
		$\widehat{proba}$	0.95	0.90	0.84	0.78
	$X^{(d,\beta)}$, $\beta=1$	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	0.071	0.068	**0.067***	0.071
		mean($\widetilde{m}_N$)	15.2	13.6	11.7	10.9
		$\widehat{proba}$	0.92	0.88	0.85	0.80

	Model	Estimates	$p=5$	$p=10$	$p=15$	$p=20$
$N=10^5$	fGn ($H=d+1/2$)	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	0.012	0.008	0.007	**0.006***
		mean($\widetilde{m}_N$)	14.0	9.8	6.9	7.9
		$\widehat{proba}$	0.92	0.90	0.87	0.85
	FARIMA$(0,d,0)$	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	0.021	**0.019***	**0.019***	**0.019***
		mean($\widetilde{m}_N$)	15.8	12.7	11.1	9.8
		$\widehat{proba}$	0.96	0.94	0.92	0.89
	FARIMA$(1,d,1)$	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	0.039	0.037	**0.035***	**0.035***
		mean($\widetilde{m}_N$)	25.7	21.8	21.4	20.4
		$\widehat{proba}$	0.98	0.98	0.94	0.93
	$X^{(d,\beta)}$, $\beta=1$	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	0.042	0.042	**0.040***	0.041
		mean($\widetilde{m}_N$)	22.3	19.9	19.7	16.9
		$\widehat{proba}$	0.99	0.97	0.93	0.90

Table 1: $\sqrt{MSE}$ *of the estimator* $\widetilde{d}_N^{(IR)}$*, sample mean of the estimator* $\widetilde{m}_N$ *and sample frequency that* $\widehat{T_N} \leq q_{\chi^2(p-1)}(0.95)$ *following p from simulations of the different processes of the benchmark. For each value of N (*10^3*,* 10^4 *and* 10^5*), of d (*-0.4*,* -0.2*,* 0*,* 0.2 *and* 0.4*) and p (*5*,* 10*,* 15*,* 20*),* 100 *independent samples of each process are generated. The values* $\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$*, mean(*$\widetilde{m}_N$*) and* $\widehat{proba}$ *are obtained from sample mean on the different values of d.*

$N = 10^3 \longrightarrow$

Model	$\sqrt{MSE}$	$d=-0.4$	$d=-0.2$	$d=0$	$d=0.2$	$d=0.4$
fGn ($H = d + 1/2$)	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.102	0.088	**0.094 ***	0.095	0.098
	$\sqrt{MSE}\ \widehat{d}_{R}$	0.091	0.108	0.106	0.117	0.090
	$\sqrt{MSE}\ \widehat{d}_{W}$	0.215	0.103	0.078	**0.073***	**0.061***
	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	**0.074***	**0.087***	0.102	0.084	0.110
	$\sqrt{MSE}\ \widehat{d}_N(10)$	0.096	0.135	0.154	0.158	0.154
	$\sqrt{MSE}\ \widehat{d}_N(30)$	0.112	0.192	0.246	0.270	0.252
FARIMA$(0, d, 0)$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.096	0.096	0.098	0.096	0.093
	$\sqrt{MSE}\ \widehat{d}_{R}$	0.094	0.113	0.107	0.112	0.084
	$\sqrt{MSE}\ \widehat{d}_{W}$	**0.069***	**0.073***	**0.074***	**0.082***	**0.085***
	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	0.116	0.085	0.103	0.094	0.101
	$\sqrt{MSE}\ \widehat{d}_N(10)$	0.139	0.133	0.148	0.146	0.156
	$\sqrt{MSE}\ \widehat{d}_N(30)$	0.157	0.209	0.232	0.247	0.243
FARIMA$(1, d, 1)$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.098	**0.092***	**0.089***	**0.088***	0.094
	$\sqrt{MSE}\ \widehat{d}_{R}$	**0.093***	0.110	0.115	0.110	**0.089***
	$\sqrt{MSE}\ \widehat{d}_{W}$	0.108	0.120	0.113	0.117	0.095
	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	0.153	0.131	0.135	0.138	0.123
	$\sqrt{MSE}\ \widehat{d}_N(10)$	0.212	0.188	0.173	0.157	0.155
	$\sqrt{MSE}\ \widehat{d}_N(30)$	0.197	0.228	0.250	0.265	0.280
$X^{(D,D')}$, $D' = 1$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.092	**0.089***	**0.113***	**0.107***	**0.100***
	$\sqrt{MSE}\ \widehat{d}_{R}$	0.093	0.111	0.129	0.124	0.111
	$\sqrt{MSE}\ \widehat{d}_{W}$	0.217	0.209	0.211	0.201	0.189
	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	**0.075***	0.101	0.121	0.122	0.131
	$\sqrt{MSE}\ \widehat{d}_N(10)$	0.109	0.143	0.163	0.168	0.180
	$\sqrt{MSE}\ \widehat{d}_N(30)$	0.109	0.177	0.228	0.249	0.247

$N = 10^4 \longrightarrow$

Model	$\sqrt{MSE}$	$d=-0.4$	$d=-0.2$	$d=0$	$d=0.2$	$d=0.4$
fGn ($H = d + 1/2$)	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.040	0.031	0.032	0.035	0.035
	$\sqrt{MSE}\ \widehat{d}_{R}$	0.040	0.027	0.029	0.031	0.030
	$\sqrt{MSE}\ \widehat{d}_{W}$	0.129	0.045	0.026	0.022	0.020
	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	**0.019***	**0.019***	**0.017***	**0.016***	**0.019***
	$\sqrt{MSE}\ \widehat{d}_N(10)$	0.036	0.038	0.049	0.043	0.048
	$\sqrt{MSE}\ \widehat{d}_N(30)$	0.043	0.070	0.086	0.081	0.076
FARIMA$(0, d, 0)$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.036	0.030	0.031	0.035	0.032
	$\sqrt{MSE}\ \widehat{d}_{R}$	0.031	0.028	0.027	0.029	0.029
	$\sqrt{MSE}\ \widehat{d}_{W}$	**0.020***	**0.018***	0.023	0.025	**0.028***
	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	0.066	0.031	**0.018***	**0.020***	**0.028***
	$\sqrt{MSE}\ \widehat{d}_N(10)$	0.076	0.047	0.043	0.053	0.038
	$\sqrt{MSE}\ \widehat{d}_N(30)$	0.074	0.085	0.073	0.086	0.073
FARIMA$(1, d, 1)$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.035	0.033	0.032	0.036	0.031
	$\sqrt{MSE}\ \widehat{d}_{R}$	**0.031***	**0.029***	**0.030***	**0.032***	**0.027***
	$\sqrt{MSE}\ \widehat{d}_{W}$	0.054	0.054	0.050	0.052	0.048
	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	0.099	0.066	0.052	0.047	0.046
	$\sqrt{MSE}\ \widehat{d}_N(10)$	0.141	0.095	0.075	0.055	0.051
	$\sqrt{MSE}\ \widehat{d}_N(30)$	0.111	0.085	0.094	0.090	0.074
$X^{(D,D')}$, $D' = 1$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.029	**0.037***	**0.035***	**0.041***	**0.038***
	$\sqrt{MSE}\ \widehat{d}_{R}$	0.032	0.041	0.037	**0.041***	0.039
	$\sqrt{MSE}\ \widehat{d}_{W}$	0.110	0.115	0.115	0.112	0.114
	$\sqrt{MSE}\ \widetilde{d}_N^{(IR)}$	**0.018***	0.064	0.092	0.084	0.081
	$\sqrt{MSE}\ \widehat{d}_N(10)$	0.035	0.093	0.102	0.106	0.094
	$\sqrt{MSE}\ \widehat{d}_N(30)$	0.039	0.088	0.084	0.074	0.077

Table 2: *Comparison of the different log-memory parameter estimators for processes of the benchmark. For each process and value of d and N, $\sqrt{MSE}$ are computed from 100 independent generated samples.*

$N = 10^3 \longrightarrow$

Model+Innovation	$\sqrt{MSE}$	$d=-0.4$	$d=-0.2$	$d=0$	$d=0.2$	$d=0.4$
FARIMA$(0,d,0)$ Uniform	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.189	0.090	0.091	**0.082***	0.092
	$\sqrt{MSE}\ \widehat{d}_{R}$	0.171	0.104	0.109	0.102	**0.086***
	$\sqrt{MSE}\ \widehat{d}_{W}$	**0.111***	**0.066***	**0.072***	0.118	0.129
	$\sqrt{MSE}\ \widetilde{d}_{N}^{(IR)}$	0.186	0.081	0.083	0.112	0.093
FARIMA$(0,d,0)$ Burr ($\alpha=2$)	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.174	0.087	0.092	0.084	**0.091***
	$\sqrt{MSE}\ \widehat{d}_{R}$	0.183	0.104	0.097	0.107	0.079
	$\sqrt{MSE}\ \widehat{d}_{W}$	**0.149***	**0.086***	0.130	0.101	0.129
	$\sqrt{MSE}\ \widetilde{d}_{N}^{(IR)}$	0.221	0.119	**0.076***	**0.082***	0.139
FARIMA$(0,d,0)$ Burr ($\alpha=3/2$)	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.188	**0.087***	**0.063***	**0.099***	0.075
	$\sqrt{MSE}\ \widehat{d}_{R}$	**0.183***	0.110	0.079	0.125	**0.072***
	$\sqrt{MSE}\ \widehat{d}_{W}$	0.219	0.108	0.138	0.146	0.159
	$\sqrt{MSE}\ \widetilde{d}_{N}^{(IR)}$	0.264	0.134	0.094	0.155	0.187
GARMA$(0,d,0)$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.149	0.109	0.086	0.130	0.172
	$\sqrt{MSE}\ \widehat{d}_{R}$	**0.098***	0.104	0.090	0.132	**0.125***
	$\sqrt{MSE}\ \widehat{d}_{W}$	0.117	**0.074***	**0.081***	0.182	0.314
	$\sqrt{MSE}\ \widetilde{d}_{N}^{(IR)}$	0.124	0.121	0.110	**0.102***	0.331
Trend	$\sqrt{MSE}\ \widehat{d}_{MS}$	1.307	0.891	0.538	0.290	0.150
	$\sqrt{MSE}\ \widehat{d}_{R}$	0.900	0.700	0.498	0.275	0.087
	$\sqrt{MSE}\ \widehat{d}_{W}$	**0.222***	**0.103***	0.083	0.071	**0.059***
	$\sqrt{MSE}\ \widetilde{d}_{N}^{(IR)}$	1.65	0.223	**0.079***	**0.050***	0.076
Trend + Seasonality	$\sqrt{MSE}\ \widehat{d}_{MS}$	1.178	0.803	0.477	0.238	0.123
	$\sqrt{MSE}\ \widehat{d}_{R}$	0.900	0.700	0.498	0.284	**0.091***
	$\sqrt{MSE}\ \widehat{d}_{W}$	**0.628***	**0.407***	0.318	0.274	0.283
	$\sqrt{MSE}\ \widetilde{d}_{N}^{(IR)}$	1.54	1.01	**0.311***	**0.158***	0.145

$N = 10^4 \longrightarrow$

Model+Innovation	$\sqrt{MSE}$	$d=-0.4$	$d=-0.2$	$d=0$	$d=0.2$	$d=0.4$
FARIMA$(0,d,0)$ Uniform	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.177	0.039	0.033	0.034	0.034
	$\sqrt{MSE}\ \widehat{d}_{R}$	0.171	0.032	0.030	0.028	**0.032***
	$\sqrt{MSE}\ \widehat{d}_{W}$	**0.125***	**0.027***	0.025	0.028	0.035
	$\sqrt{MSE}\ \widetilde{d}_{N}^{(IR)}$	0.165	0.042	**0.017***	**0.027***	**0.032***
FARIMA$(0,d,0)$ Burr ($\alpha=2$)	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.180	0.036	0.041	0.033	0.032
	$\sqrt{MSE}\ \widehat{d}_{R}$	0.169	**0.031***	0.030	**0.031***	**0.029***
	$\sqrt{MSE}\ \widehat{d}_{W}$	**0.138***	0.068	0.065	0.076	0.066
	$\sqrt{MSE}\ \widetilde{d}_{N}^{(IR)}$	0.219	0.067	**0.018***	0.039	0.074
FARIMA$(0,d,0)$ Burr ($\alpha=3/2$)	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.18	0.038	**0.026***	0.030	**0.021***
	$\sqrt{MSE}\ \widehat{d}_{R}$	0.174	**0.033***	0.031	**0.023***	0.023
	$\sqrt{MSE}\ \widehat{d}_{W}$	**0.126***	0.058	0.149	0.124	0.090
	$\sqrt{MSE}\ \widetilde{d}_{N}^{(IR)}$	0.264	0.113	0.030	0.099	0.159
GARMA$(0,d,0)$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.063	0.041	0.028	0.032	0.060
	$\sqrt{MSE}\ \widehat{d}_{R}$	**0.037***	**0.033***	0.025	**0.026***	**0.030***
	$\sqrt{MSE}\ \widehat{d}_{W}$	0.061	0.052	0.021	0.078	0.081
	$\sqrt{MSE}\ \widetilde{d}_{N}^{(IR)}$	0.074	0.040	**0.016***	0.055	0.109
Trend	$\sqrt{MSE}\ \widehat{d}_{MS}$	1.16	0.785	0.450	0.171	0.072
	$\sqrt{MSE}\ \widehat{d}_{R}$	0.900	0.700	0.431	0.192	0.067
	$\sqrt{MSE}\ \widehat{d}_{W}$	0.135	0.046	**0.021***	0.019	0.021
	$\sqrt{MSE}\ \widetilde{d}_{N}^{(IR)}$	**0.019***	**0.021***	**0.021***	**0.016***	**0.020***
Trend + Seasonality	$\sqrt{MSE}\ \widehat{d}_{MS}$	1.219	0.841	0.474	0.194	0.099
	$\sqrt{MSE}\ \widehat{d}_{R}$	0.900	0.700	0.431	0.189	0.063
	$\sqrt{MSE}\ \widehat{d}_{W}$	**0.097***	**0.073***	0.063	0.065	0.051
	$\sqrt{MSE}\ \widetilde{d}_{N}^{(IR)}$	0.671	0.382	**0.049***	**0.047***	**0.041***

Table 3: *Comparison of the different log-memory parameter estimators for processes of the benchmark. For each process and value of d and N, $\sqrt{MSE}$ are computed from 100 independent generated samples.*